现代数学基础丛书

神经动力学模型方法和应用

● 阮 炯 顾凡及 蔡志杰 编著







科学出版社

(O-1279.0101)

责任编辑：吕 虹

封面设计：槐寿明



现代数学基础丛书

巴拿赫空间引论
近代调和分析方法及其应用
递归论
拓扑群引论
公理集合论导引
丢番图逼近引论
Banach代数
紧黎曼曲面引论
广义哈密顿系统理论及其应用
解析数论基础
数理统计引论
多元统计分析引论
概率论基础
微分动力系统原理
二阶椭圆型方程与椭圆型方程组
分析概率论
非线性发展方程
黎曼曲面
傅里叶积分算子理论及其应用
反应扩散方法引论
离散鞅及其应用
复合算子理论
微分方程定性理论
概率论基础和随机过程
复解析动力系统
模型论基础
环与代数
仿微分算子引论
实分析导论
对称性分岔理论基础
线性微分方程的非线性扰动
随机点过程及其应用
复变函数逼近论
线性整数规划的数学基础
组合矩阵论
算子代数
Banach空间中的非线性逼近理论
Gel'fond-Baker方法在丢番图方程中的应用
实用微分几何引论
半群的S-系理论
有限典型群子空间轨道生成的格
有限群导引（上册、下册）
随机模型的密度演化方法
非线性偏微分复方程
调和分析及其在偏微分方程中的应用
惯性流形与近似惯性流形
数学规划导论
拓扑空间中的反例
拓扑空间论
非经典数理逻辑与近似推理
光滑映射的奇点理论
神经动力学模型方法和应用



ISBN 7-03-008782-8/O·1279
定 价：28.00 元

内 容 简 介

神经网络动力学是神经网络与动力学系统交叉结合的学科。本书着重阐述神经网络动力学模型的建立和各类模型研究的问题与方法。书中还介绍了神经动力学的生物背景与相关的动力系统理论的基础，并且给出了国内外一些应用的成果介绍。

读者对象为大专院校理工科各专业的大学生、研究生、教师及有关的科技工作者。

图书在版编目(CIP)数据

神经动力学模型方法和应用/阮炯，顾凡及，蔡志杰编著．—北京：科学出版社，2002

(现代数学基础丛书)

ISBN 7-03-008782-8

Ⅰ．神…　Ⅱ．①阮…②顾…③蔡…　Ⅲ．神经网络-动力学-生物模型-研究　Ⅳ．Q811.2

中国版本图书馆 CIP 数据核字(2000)第 44338 号

科学出版社 出版

北京东黄城根北街16号

邮政编码：100717

http://www.sciencep.com

源海印刷厂 印刷

科学出版社发行　各地新华书店经销

*

2002 年 4 月第 一 版　开本：850×1168　1/32

2002 年 4 月第一次印刷　印张：10 7/8

印数：1—2 000　字数：286 000

定价：28.00 元

(如有印装质量问题，我社负责调换〈北燕〉)

前　　言

神经网络在最近 20 年中得到了迅速的发展，有关的论文及著作已有许多．本书并不想全面介绍神经网络，而只是想对神经网络与数学中动力学系统理论的交叉结合介绍一些我们在学习与研究中的体会．

我们力图在以下几个方面作一些尝试：(1) 将人工神经网络模型与生物神经网络模型用动力学观点统一起来，突出神经动力学的思想；(2) 把吸引子 (包括收敛的、振荡的、混沌的三种类型) 的研究贯穿在神经网络的理论与应用的整个过程之中；(3) 使应用数学工作者，通过本书可以了解动力系统中吸引子分析在神经网络研究中的对应问题与应用，而使从事生物、计算机、信息等学科研究的工作者，通过本书可以了解神经网络如何作为动力学系统来认识，怎样以吸引子分析的主线贯穿整个神经网络研究之中．

本书分四章．第一章讲述神经动力学的生物学基础；第二章介绍动力系统理论的基础知识；第三章是本书重点，介绍各类神经网络的动力学模型与问题，并讲述分析的方法与思想；第四章给出一些神经动力学应用的例子．

全书主要介绍国内外学者近几十年来在神经网络动力学方面的研究成果，其中包含了作者以及复旦大学神经网络研究小组同仁们的一些研究工作．

全书第一章由顾凡及执笔；第二章由阮炯执笔；第三章 §1～§6 及 §10 由阮炯执笔，§7～§9 由顾凡及执笔；第四章由蔡志杰执笔．在整个编写过程中林伟、黎丽娟两位同志帮助作者收集资料，并打印了大部分的书稿，在此特别向他们致谢．在本书的编写、出版过程中得到了科学出版社吕虹同志的大力支持与帮助，在此表示衷心感谢．

由于我们水平有限，在书中难免有错误与不妥之处，恳请读者批评指正.

编　者

2001.12

目 录

第一章　神经元的神经生物学背景

神经细胞或者叫做神经元是神经网络的结构单元，也是它的功能单元. 由于本书的主题是神经网络，因此我们是在这一层次上讨论问题，从而只把神经元作为构成神经网络的基本元件，而不把它本身也当作一个由更下一层次的元件 —— 细胞膜上的膜电容、膜电导和各种离子通道等所构成的系统. 这样，在这一章里我们只介绍神经元从信息处理的角度来看最为重要的外特征，而不去探究造成这些特性的内部机制.

§1.1　神经元的结构

尽管我们在上面声称神经元是神经系统的基本单元，但这决不意味着神经系统 (即使是同一个动物中的神经系统) 中的所有神经元都是一模一样的，好像由砖块构造的砖房中的砖块一样. 事实上，神经系统中神经元的形态是多种多样的. 粗粗分一分，在哺乳动物的脑中可区分出几何形状彼此不同的神经元多达 200 种以上，图 1.1.1 中给出的仅仅是其中的几个例子. 这些细胞的形状和大小各异，反映了它们在处理信息的方式上也各有不同. 应该牢记的是神经元决不仅仅是传输信息的中继站，典型的神经元总是从若干来源接受信号的，把这些信号整合起来再编码成复杂的输出信号，然后再把这一信号传输到许多别的细胞上去.

尽管神经元的这种多样性使问题复杂化，但是对许多神经元而言，我们还是可以把它们简化成由如图 1.1.2 所示的 3 个部分组成：细胞体、树突和轴突.

(1) 细胞体　神经元如同其它细胞一样有一个含有细胞核的细胞体，在其中进行复杂的代谢活动. 在细胞体的膜上有许多特异化

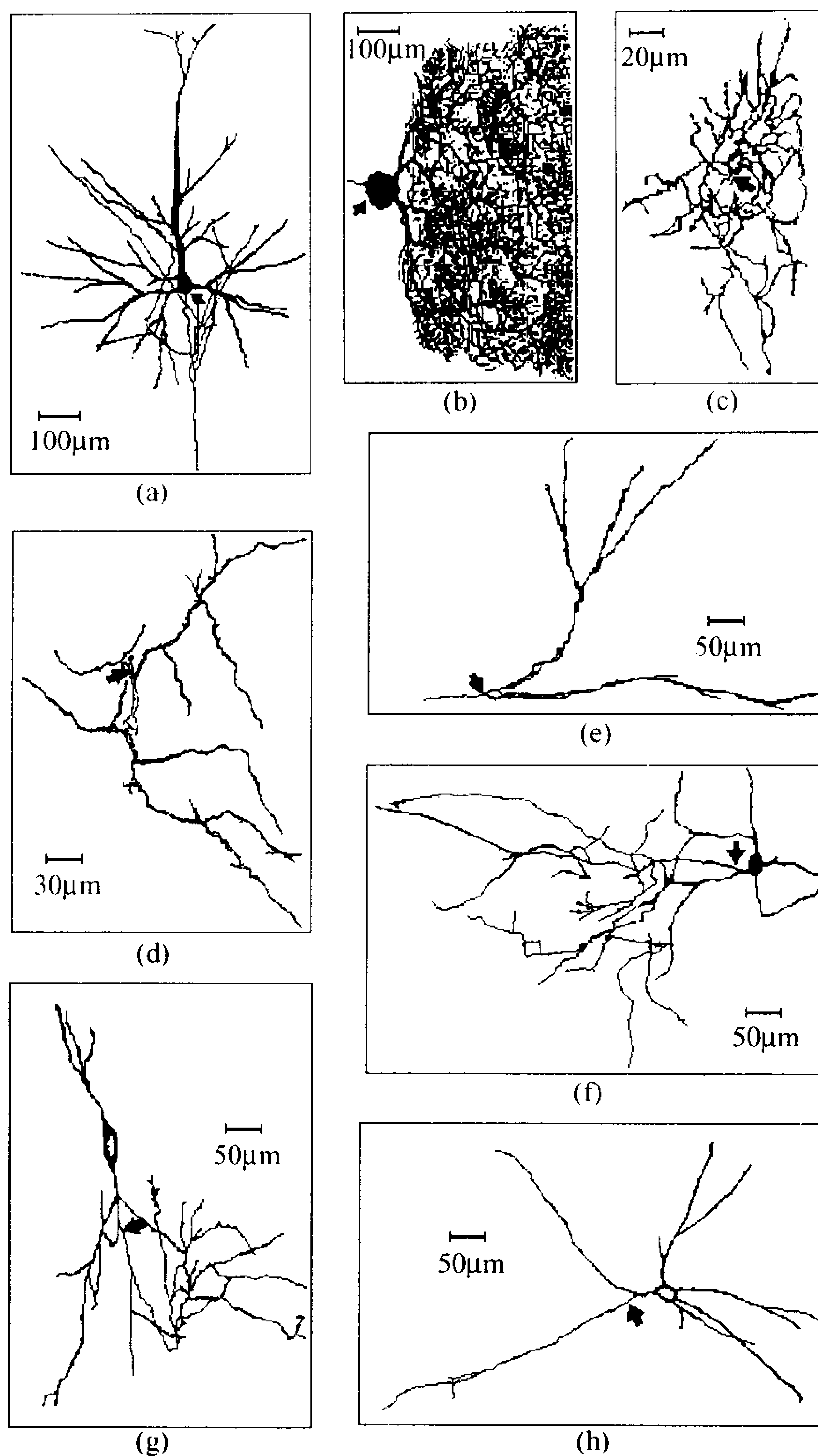
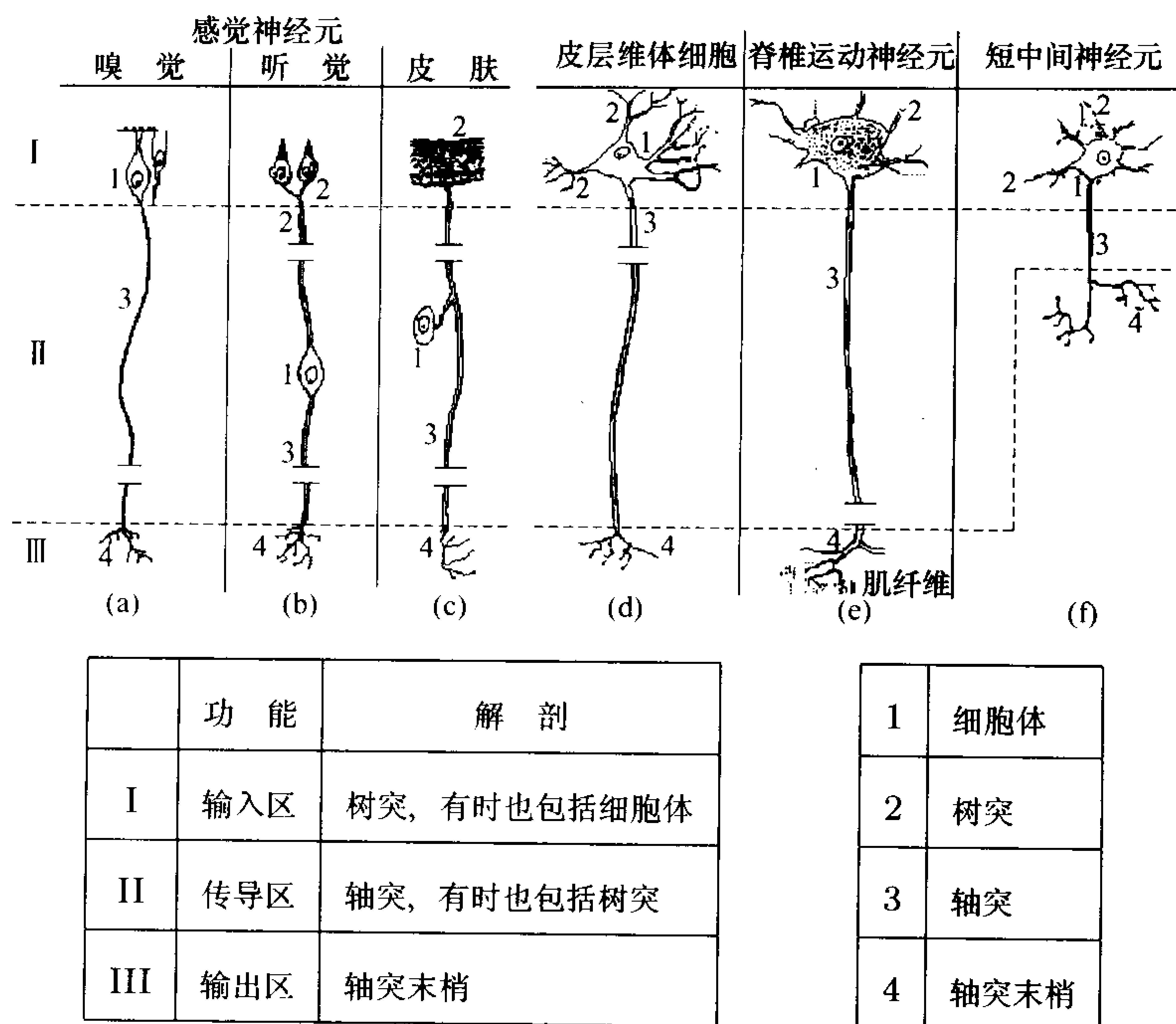


图 1.1.1　某些不同类型的神经元

(a) 大脑皮层中的锥体细胞;　(b) 小脑皮层中的浦金野氏细胞;　(c) 小脑皮层中的篮状细胞;　(d) 桥脑中的多极细胞;　(e~h) 在海马中可辨认出的 21 种细胞的四种. 细胞中轴突起始处以箭头指出. 另一些类型的神经元见图 1.1.2

(引自 Rosenzweig and Leiman, 1982)



	功 能	解 剖
I	输入区	树突，有时也包括细胞体
II	传导区	轴突，有时也包括树突
III	输出区	轴突末梢

1	细胞体
2	树突
3	轴突
4	轴突末梢

图 1.1.2 神经元中的各个结构和功能部分

(引自 Rosenzweig and Leiman, 1982)

的小区域，其中一些是对特异性的化学递质起作用的受体，另一些则从外界吸收营养或向外界排除废物．但是从信息处理的观点来看，细胞体和树突是专门用来接收和整合信息的，而轴突则是用以把信息传递给其它细胞．

(2) 轴突 如图 1.1.1 和图 1.1.2 所示，神经元在一端向外发出一个很长的突起，其长度因细胞而异，短的只有几微米，长的可以达 1 米以上．对许多神经元来说，轴突在其离开细胞体之外形成一个锥形区域称为**轴丘**．除了极少的例外，神经元都只有一个轴突，然而从一根轴突上可以分出许多侧枝，从而对许多其它神经元产生影响．在轴突或其侧枝的末端常常分成许多直径很细的末梢，它们与其它细胞的细胞体或树突形成联系．

(3) 树突　如图 1.1.1 和图 1.1.2 所示，神经元除了在其一端向外发出轴突之外，大多数神经元的细胞体还向其它各个方向发出许多突起，这些突起又以非常复杂的形式分出许多分枝. 各种不同类型的神经元其主要区别就在于这些像树枝状的突起 —— 树突形式各不相同. 这些形状复杂的树突大大扩大了神经细胞的表面积，使它们可以接收更多来自其它神经元的输入. 现在已经知道，像皮层锥体细胞这样的神经元，其树突部分占到细胞总体体积的 95%.

(4) 突触　一般来说，轴突的末梢端往往形成某种膨起，称为**终扣**. 在终扣中有许多小泡，其中含有一定量的化学物质 —— 神经递质. 终扣并不和其它的神经元直接连在一起，它终止在其它细胞的细胞体或树突的邻近，中间隔开一个 20 ~ 40nm 的空隙，称为 **突触间隙**. 在突触间隙后面的下一个神经元的细胞膜 —— 突触后膜上具有对某些神经递质能起特异化反应的蛋白质 —— 受体. 当前一个神经元终扣释放出某种神经递质，通过突触间隙扩散到达另一个神经元的突触后膜时，和相应的受体结合而打开某种离子通道. 终扣末端的细胞膜 —— 突触前膜、突触间隙和突触后膜这三者合在一起完成信息从前一个细胞向后一个细胞的传递，这三者的总体就称为 **突触** (见图 1.1.3).

大多数突触发生在树突上. 在不少情况下，树突上还会长出一个有小苞的圆盘状 (称为 **侧棘**), 并在此处形成突触 (见图 1.1.3), 在细胞体上形成的突触只占突触总数中的一小部分.

在树突上形成的突触称为 **轴突突触**, 而在细胞体上形成的突触则称为 **轴体突触**. 此外，在某些情况下，树突和树突之间，或者轴突和轴突之间也会形成突触.

上述突触都是通过化学递质的释放、扩散和结合来实现信息从一个神经元向下一个神经元的传递，因此称为 **化学突触**, 这种突触在高等动物的神经系统中占了绝大多数. 但也还有另一类突触称为 **电突触**, 这类突触的间隙要比化学突触的小得多，离子和小分子可直接从一个细胞流入其相邻的细胞. 离子从一个细胞通过通

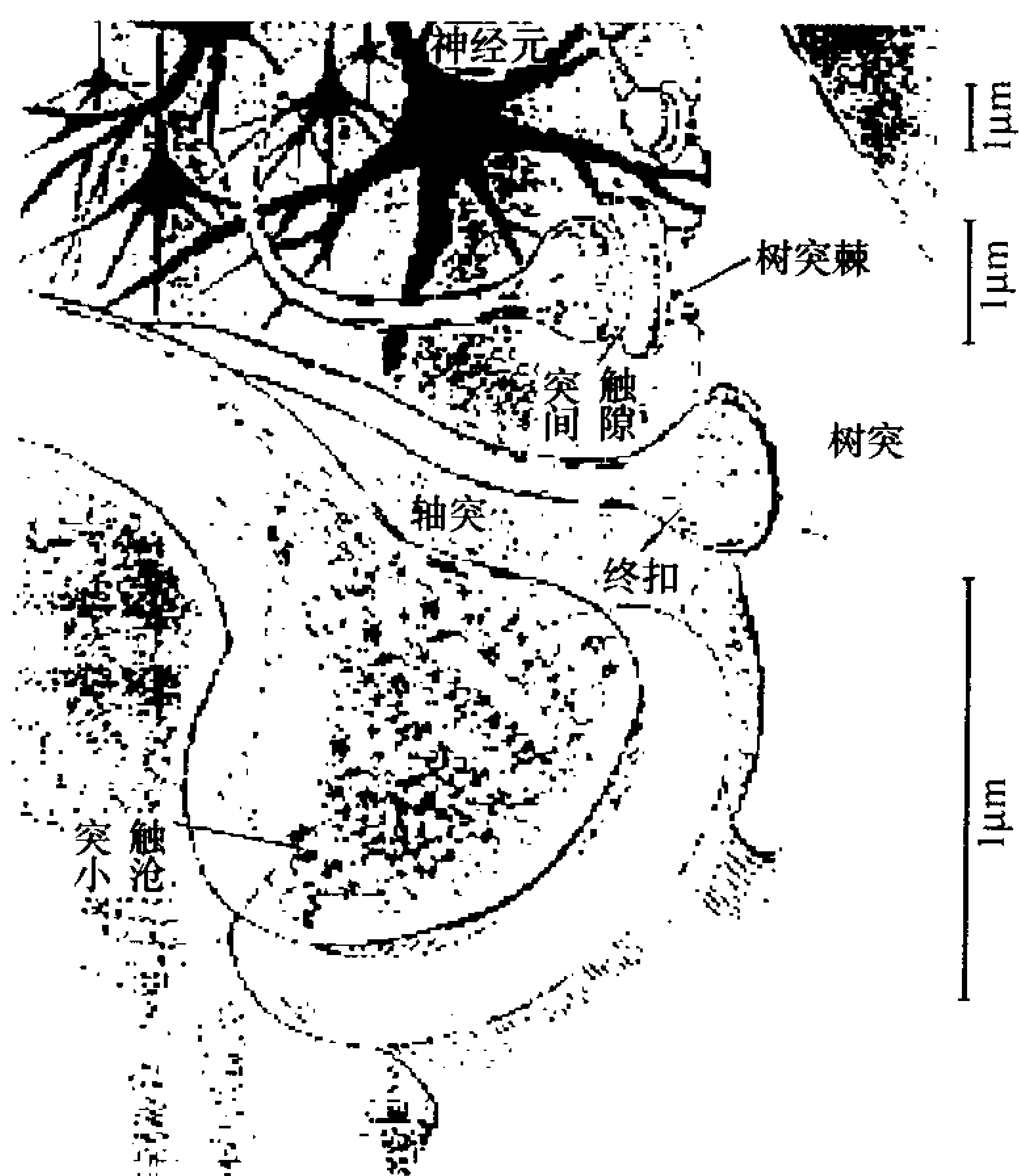


图 1.1.3　在树突表面和侧棘处形成的突触

(引自 Rosenzweig and Leiman, 1982)

道流入另一个细胞就可使一个细胞的电变化几乎即时传递到另一个细胞. 在大多数情况下, 离子流可以是双向的, 但是某些电突触离子流只能沿单向流动, 这种突触就称为 **整流突触**.

§1.2　神经元的功能特征

和神经元的结构相对应, 从功能上来看神经元大致上可以分成 3 个部分: 树突和细胞体上的突触是它的接收端或输入端, 细胞体特别是轴丘以及紧接轴丘的一小段裸露在细胞外液中的轴突 (这两者合在一起称为 **始段**) 是整合来自各个突触的信息的场所, 而轴突则是它的输出端. 图 1.2.1 给出了神经元的这种功能分区的一个示意图.

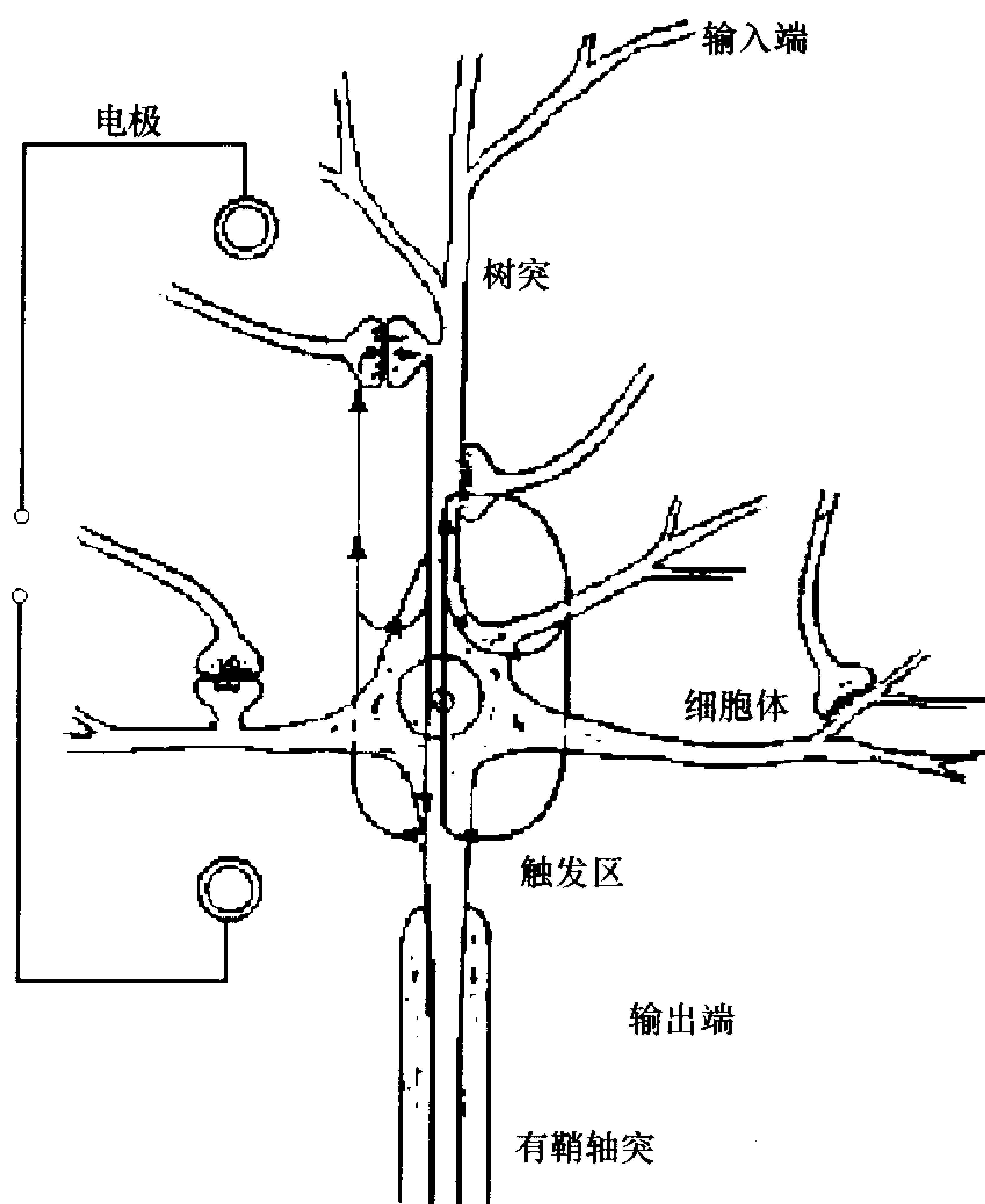


图 1.2.1　神经元的功能组织模式图

(引自 Freeman,1992)

通常一个神经元和其它神经元要形成上千个突触联系. 当突触前膜有神经冲动到达时, 终扣中的突触小泡会向突触间隙中释放所含的神经递质, 这些神经递质通过突触间隙扩散到突触后膜, 并和其上的受体结合, 而使离子通道开放, 从而使某些离子流涌入后一个神经元或自其中流出. 前者将使细胞膜内外的电位差减小, 称为 **去极化**, 其结果是神经元趋向兴奋; 而后者则使这种电位差进一步增大, 称为 **超极化**, 其结果是神经元趋向抑制. 究竟其结果属于何种情况则取决于神经递质的性质. 由上所述可知突触可以分为两类: 兴奋性突触和抑制性突触. 由于神经递质的释放、扩散以及受体的结合导致离子通道的开放都有一个时间过程, 因此从前一个神经元的神经冲动到达其神经末梢到后一个神经元的突触后

电位 (PSP) 发生变化有一段时间延迟，称为 **突触延搁**.

根据 Kirchhoff 定律流入或流出突触的电流都必须形成一个回路，因此当电流以某一方向流经突触时，都要以相反方向流经始段膜，这是由于细胞体上只有这一区域是裸露在细胞外液中，阻抗最小，而细胞体的其它表面都为其它神经元神经末梢的终扣以及另一种细胞 —— 神经胶质细胞的组织所覆盖，具有很高的阻抗. 因此虽然电流从激活的突触流经始段的过程中也会在沿路通过细胞体的膜有所分流，但这种分流是很小的. 这样，来自各个突触的电流在流经始段膜时就会总和起来，称为 **空间总和**. 从功能的角度来说，我们可以把始端段看成是将来自各个突触的信息进行整合的区域. 由于来自兴奋性突触的电流和来自抑制性突触的电流在始段处方向正好相反，因此这种总和是代数和. 这一总电流在流经始段膜的电阻时形成某种电压降. 当这一电压降达到某种数值时就会在始段处触发一个神经冲动，这一电压降的值称为 **阈值**. 由于神经冲动是在始段处触发产生的，因此我们也常把始段处称为神经元的 **触发区**.

神经冲动也称为 **动作电位**. 它有一个特点，即在正常生理状态下，一旦膜内外的电压差达到阈值触发产生一个动作电位，这个电位的幅度和波形都是固定不变的，而且当它沿轴突传播时，其波形和幅度也并不发生变化，因此动作电位要么不被触发产生，而一旦被触发产生其幅度和波形就是固定的. 动作电位的这一性质被称为动作电位的 **全或无定律**. 动作电位的持续时间约为 1 毫秒. 动作电位的全或无定律保证了它可以把它所携带的信息不失真地沿轴突全线进行传输.

动作电位一旦产生，在其后一段时间内阈值上升到无穷大，亦即在这段时间里，不论神经元接收到多强的输入都不能再产生另一个动作电位，这一段时间就称为 **绝对不应期**. 在经过绝对不应期之后，阈值近似地按指数下降逐渐恢复到正常值，亦即在这一段时间内，只有当输入超过正常强度时才能触发起另一个动作电位，这段时间就称为 **相对不应期**. 如果观察得很仔细的话，还可以

发现从相对不应期过渡到正常阈值时还要经过一小段时间，在这段时间里阈值甚至低于正常值，这段时间被称为 **超常期**.

如果我们要仔细分析神经细胞膜上的电位变化过程，我们必须对膜的性质加以观察. 由于细胞膜不允许离子自由通过，因此在细胞内 (膜内) 和细胞外 (膜外) 的各种离子浓度是不一样的，所以细胞膜起着某种电容的作用. 当然细胞膜也并不是绝对绝缘的，即使在不被激发的状态下，它也允许有些离子通过膜形成跨膜的漏电阻. 因此一般说来可以把细胞膜想象成由电阻和电容并联而成的等效电路. 由于存在电容，当从某一神经纤维向某一突触传来一串神经冲动时，如果这些神经冲动之间的间隔足够短，那么当前一个脉冲在突触后膜上引起的电位变化还未全部消退之前，后一个脉冲又将在此基础上引起新的电位变化而叠加在一起，这被称为 **时间总和**.

如果我们通过微电极向神经元注入一定强度的电流并使此电流强度保持不变而持续一段时间，同时测量神经元所发放的脉冲序列在经过一段过渡过程而最后达到稳态时的发放率 (单位时间内的脉冲数)，我们就可以发现这个发放率与注入电流强度之间服从一种 S 形曲线的函数关系 (见图 1.2.2). 这一关系被称为神经元的 **输出函数**.

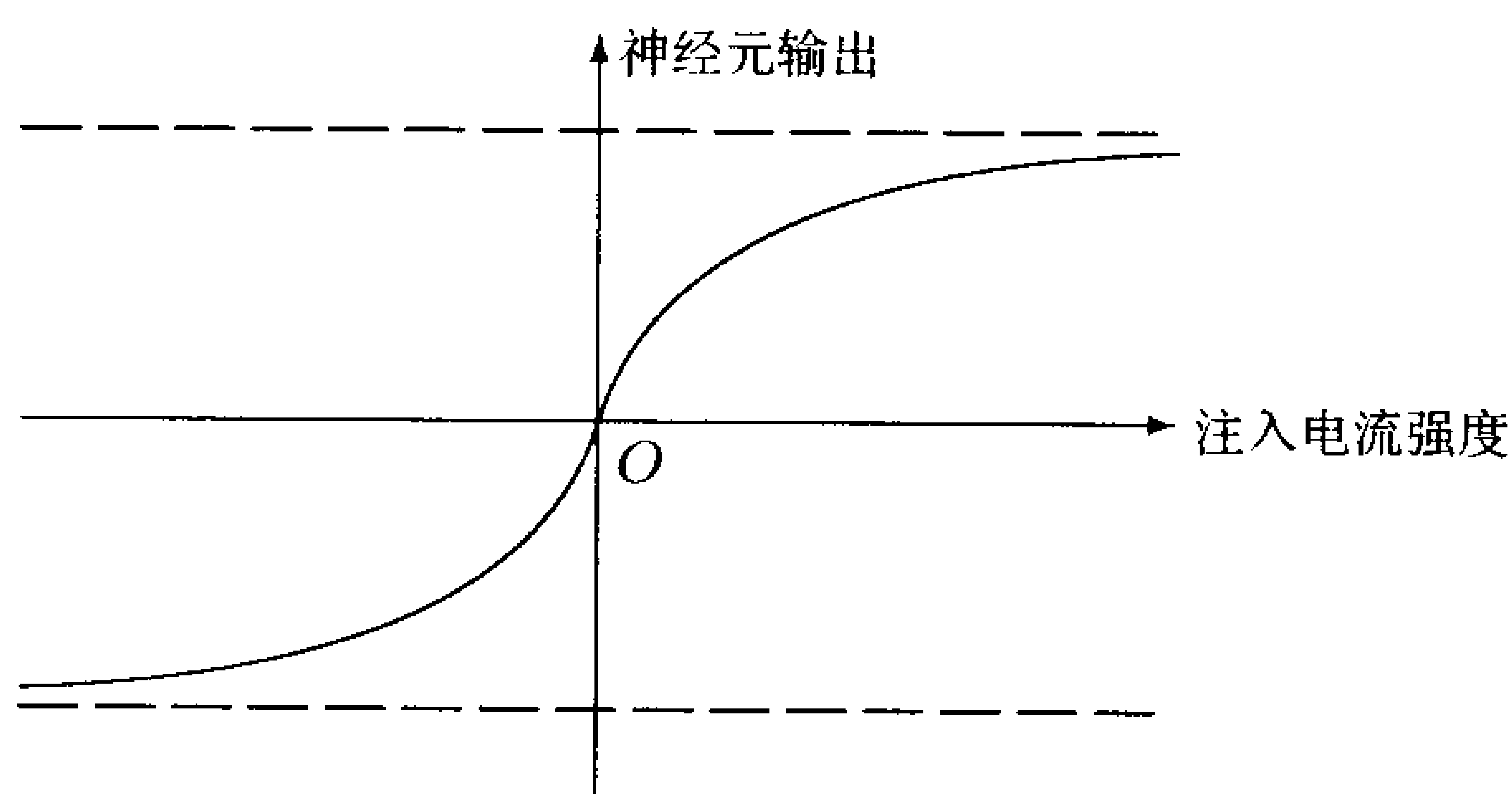


图 1.2.2 神经元的输出函数

此外，对于许多神经元来说，当它刚接收输入时，其脉冲发放

率较密，如果输入强度维持不变，这一发放率将会逐渐下降，而最后稳定在某一水平，这种现象称之为 **适应**.

最后，我们还要指出在神经元发放一个脉冲之后，膜电位重新恢复到原来未被激活时的水平 (静息电位)，这被称为 **复极化**，甚或比静息电位更负的水平 (在不被激活时细胞内的电位比细胞外要负. 当细胞接收兴奋性刺激时，这一电位差减小，亦即细胞内电位向细胞外电位靠近，即去极化. 当发生动作电位时，其峰值甚至可使细胞内电位较膜外为正)，这被称为 **超极化**.

总结起来说，从信息处理的观点来看，大多数神经元主要具有下列外特性:

(1) 突触有两类：兴奋性突触和抑制性突触;

(2) 存在突触延搁;

(3) 神经细胞有 RC 性质，并因此具有时间总和作用;

(4) 存在空间总和作用;

(5) 存在阈值;

(6) 脉冲发放后发生复极化或超极化;

(7) 发放脉冲后有一定的不应期 (绝对不应期和相对不应期);

(8) 动作电位服从全或无定律;

(9) 输出函数呈 S 形曲线;

(10) 存在适应性.

§1.3　人工神经网络和现实性神经网络

数字电子计算机的发明和发展改变了整个人类社会. 在不到半个世纪的时间里，计算机已经历了从电子管到晶体管到集成电路到大规模、超大规模集成电路这样的几个飞跃. 计算机的计算速度越来越快，存储容量越来越大，它已能实现许多以往只有人脑才能实现的功能，从而被人们称为电脑. 然而它从原理上和人脑却相差甚远. 首先，计算机是串行处理的，而人脑是并行处理的；其次，计算机是按地址记忆的，而人脑则是按内容记忆的；再次，计

算机的运算单元和存储单元是明确区分开来的，而人脑则没有这种明确的区分，记忆分布在整个大脑．此外，计算机特别擅长于数学运算和逻辑运算，然而对即使是幼儿的大脑也能轻而易举地做到的从众多景物中识别某一特定目标，譬如其母亲这样简单的任务，对于计算机来说也是十分困难的．计算机需要人给它编制程序才能执行其特定的功能，而人脑则可以通过经验自觉地学会本领．为了弥补计算机的不足之处，使计算机更加智能化，人机界面对人更友善，人们正试图发展新一代的仿脑计算机．这种机器是由大量像神经元这样的元件组成的，进行并行处理，其记忆是按内容分布存储的，具有自组织和学习的能力．神经网络的兴起正是顺应了这一要求．

神经网络的基本元件是神经元，亦即具有生物神经元某些特性的元件．如同我们在本章一开始声明的那样，不把神经元本身当作一个系统来处理，亦即不考虑其内部的离子通道机制 (作如上考虑的神经元模型常被称为神经元的 **生物物理模型**，这种模型固然对理解神经元的内部机制是重要的，然而对于理解神经网络的功能常常是不必要的，它给在神经网络水平上研究的问题增添了非常大的而不必要的复杂性，这正如一位计算机工程师不必去管半导体物理中的细节一样)，那么现有的模型通常都是从上节的 10 条性质中，抽取其中的某些性质来建立模型．从应用的角度来说，人们只要求神经网络能解决他们所希望解决的实际问题，而不去管神经系统中是否也是按同样的方法来解决，甚至其基本单元 —— 神经元的性质也可能不完全符合生物神经元的真实情况．因此这种神经网络被称为 **人工神经网络**，或是由生物神经网络中得到某些启发而建立的模型．

另一类神经网络其主要目的是试图以此作为理解生物神经信息处理的工具．因此，它首先要求其基本单元的特性不能违背生物神经元本身固有的特性，一般都是从上节的 10 条性质中选取若干条作为建立其神经元模型的前提条件；其次，由这些神经元构造的神经网络，在各神经元的联结方式上多采取某个真实生物神经回

路的内部结构方式，因此常可通过对此类神经网络的仿真来解释相应神经回路的信息处理机制. 这一类神经网络常被称为 **现实性神经网络**.

根据对什么是神经元的输出变量，或者说根据对神经编码的不同看法，可以把现实性神经元模型分为两大类. 一类认为神经元的输出信息是由其发放率携带的，这是一个模拟量，它和膜电流或刺激之间存在着某种单调有界不减函数的关系，我们把这一类神经元模型称为 **模拟神经元模型**. 上述性质往往是此类模型中最主要的非线性性质，因此在由它构成的神经网络的动力学性质中扮演着重要的角色. 另一类模型则认为神经元的输出是一串脉冲序列，不仅这串序列的发放率，亦即其在单位时间内的脉冲数，或者说是其一阶统计量携带有信息，而且在一定时间窗内这一序列的时间模式，或者说其高阶统计量也携带有信息. 这类模型常被称为 **脉冲发放神经元模型** (spiking neuron model) 或 **合后发放模型** (integrate–and–fire model). 从目前的情况看，大多数的神经网络都采用模拟神经元作为其基本单元. 但近年来有越来越多的人认识到发放率编码并不是普适的，在许多场合下神经元发放的脉冲序列的时间模式起着至关重要的作用，因此对脉冲方法神经元模型的研究正在增多.

我们相信，对现实性神经网络的研究将不仅有助于我们揭开脑工作原理的奥妙，而且一旦当我们真正地在这方面有所进展并将它借鉴到人工神经网络方面来，它还必将推动人们设计出更接近于人脑的智能机器来.

最后应该指出，在上几节中我们虽然论述并总结了迄今为止神经网络工作者在他们建立模型时用到的神经元信息处理的主要性质，但近几年来神经科学的飞速发展，使我们对神经元如何工作又有了不少新的认识，一些传统的观点正在被修正. Koch (1997) 在最近的一篇综述中特别提到了有关神经元如何进行计算的三方面的进展：有关树突的研究、脉冲及其发放时刻以及突触的可塑性问题. 传统上一般把树突看成某种无源电缆，然而现在已经知道

树突膜上也有多种离子通道，从而能实现像乘法之类的非线性运算；其次，现在人们认识到至少在一些情况下，例如双耳定位，神经元能分辨时间到很高的精度，因此目前关于神经元究竟是一种积分器还是一种同时检测器正引起人们的激烈争论；最后，树突的去极化也可能引起细胞体动作电位的产生，并反传到突触，从而对突触的可塑性产生影响．这些性质目前在神经网络的研究中还很少考虑，但正如一些学者所指出的那样，神经网络今后的发展必须来源于神经生物学，必须考虑神经生物学研究进展所获得的新知识.

参考文献

[1] W. J. Freeman, Tutorial on neurobiology: from single neurons to brain chaos, International Journal of Bifurcation and Chaos, 2(1992), 451～482.

[2] J. E. Dowling, Neurons and networks —— An introduction to neuroscience, The Belknap Press of Harvard University Press, Cambridge, Massachusetts, 1992.

[3] M. R. Rosenzweig & A. L. Leiman, Physiological psychology, D. C. Heath and Company, Lexington, Massachusetts, 1982.

[4] C. Koch, Computation and the single neuron, Nature, 385(1997), 207～210.

第二章　非线性动力学模型与问题，分析与方法

本章将介绍有关非线性动力学的一些基础知识. 首先讨论非线性离散动力学模型、连续动力学模型、时滞动力学模型、时空动力学模型, 然后对这些模型中提出的动力学问题吸引子、排斥子的判别及分析作了介绍.

§2.1　非线性动力学模型与问题

从自然界中千变万化的现象出发，可以用各种数学模型去描述它们, 但总体上可以分为两大类型, 一类是确定性的数学模型, 另一类是随机性的数学模型. 而在确定性的数学模型中最普遍的是动力学模型, 即系统的状态随时间演化的过程的模型. 这一节主要介绍这类模型的四种类型，通过一些实例引出这些模型的一般形式以及相关的问题.

2.1.1　一个种群生态学的实例

在种群生态学的研究中, 针对不同的情形, 分别可以用离散模型、连续模型、时滞模型、时空模型来描述种群的变化.

1. 虫口模型

考虑像蚕、蝉这种类型的昆虫数目 (即 **虫口**) 的变化，讲得确切一些，即相邻两代 (或者说今年与明年) 的虫数的变化规律 (或者说相依关系).

假定所考虑的昆虫，每年夏季成虫产卵后全部死亡，第二年春天每个虫卵孵化成一只虫子. 显然牛、马、羊以及人均不属于此列. 设第 n 年的虫口数目为 P_n, 每只成虫平均产卵 C 个，下一年，即第 $n+1$ 年的虫口数目为 P_{n+1}, 显见有如下关系式:

$$P_{n+1} = CP_n, \quad n = 0, 1, \cdots. \tag{2.1.1}$$

这是一个线性定常差分方程，其中 C 表示此类昆虫所生活的周围环境造成的综合作用的参数.

对 $C>1$ 的情形，虽然虫口数目越来越多，但实际上周围环境能提供的空间与食物都是有限的，虫子之间为了生存将互相竞争而咬斗，同时对虫子存在威胁的因素例如传染病及天敌，综合起来将会抑制昆虫无限地增加以达到一定的平衡，也就要使下一年的虫口数目减少. 由于咬斗和接触，都是发生在两只虫子之间的事件，而 P_n 只虫子配对的事件总数是 $\frac{1}{2}P_n(P_n-1)$, 当 P_n 相当大时，这总数接近于 $\frac{1}{2}P_n^2$, 因此 (2.1.1) 模型修改后得到如下的形式:

$$P_{n+1}=CP_n-bP_n^2. \tag{2.1.2}$$

在做一些变量与参数代换后可以写成标准形式:

$$x_{n+1}=\lambda x_n(1-x_n),\quad \lambda\in[0,4],\quad n=0,1,2,\cdots. \tag{2.1.3}$$

这就是近 20 多年来引起人们极大兴趣的 **Logistic 方程**.

2. 单种群连续模型

对诸如人、牛、马、羊等种群，它们的生命期长，世代重叠并且数量很大，与前面讨论的生命期短，世代不重叠，诸如蚕蝉类虫口不同，当种群总数很大时，种群的适当变化与总数之比仍是很小的，这相当于在时间作微小变化时，种群数也作小的变化. 因此，我们近似地认为，当总数很大时，种群数对时间相依的函数关系是连续的，甚至是可微的. 这时可取时间 t 为自变量，以种群数 $N(t)$ 为状态变量，考虑 $N(t)$ 随时间 t 连续演化的方程，特别地考虑它的变化率 $N'(t)$ 与 $N(t)$ 的相依关系，建立的数学模型是常微分方程.

假定 $r(t,N)$ 表示该种群出生率和死亡率之差，那么 $N(t)$ 的增长率

$$N'(t)=r(t,N(t)). \tag{2.1.4}$$

若种群是孤立的，出生率与死亡率均与种群 $N(t)$ 成比例关系，则

演化方程为

$$N'(t) = CN(t). \tag{2.1.5}$$

这是一个一阶线性常系数常微分方程，它描述了 **Malthus 生物总数增长定律**.

当时间离初始时刻越来越长的时候，人口数或种群数不能以指数形式增长, 这是由于人们或种群要受生存空间与食物的限制, 人们之间为求得生存必导致竞争，且疾病也要威胁人们与种群，因此在上述模型中, 要考虑减去由于竞争引起的项. 因为单位时间内两个成员相遇次数的统计平均值与 N^2 成正比，所以这项取为 $-bN^2$, 即微分方程将被修改为

$$\frac{\mathrm{d}N}{\mathrm{d}t} = aN - bN^2. \tag{2.1.6}$$

这个方程称为 **生物总数增长统计筹算律**, 其中常数 a 与 b 称为生物总数的 **生命系数**. 这是一个一阶非线性定常微分方程，它描述了 Verhulst 首先发现的定律.

3. 具时滞的单种群模型

对模型 (2.1.6), 如果考虑幼虫到成虫的成长期及一些动物的怀胎期因素，例如澳大利亚袋鼠，那么在 t 时刻的种群增长率与在 $t-T$ 时刻的种群数量有关， (2.1.6) 将被修改为如下的具时滞形式的 Logistic 方程:

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = rN(t)\left[1 - \frac{N(t-T)}{K}\right] \tag{2.1.7}$$

或

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = rN(t)\left[1 - \frac{1}{K}\int_{-\infty}^{t} W(t-s)N(s)\mathrm{d}s\right]. \tag{2.1.8}$$

这种类型的数学模型称为 **具时滞的泛函微分方程** 或简称为 **时滞微分方程**.

这类模型在具有潜伏期的流行病传播的问题中很普遍. 例如 1973 年， W. P. London 和 J. A. Yorke 研究的麻疹传播的模型为

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = \beta(t)S(t)[S(t-12) - S(t-14) - 2r] + r, \tag{2.1.9}$$

这里的时滞量 12, 14 是相当大的数量，而在前面模型 (2.1.7) 中，在有的情形下，时滞量 T 是很小的甚至接近于零，这也就是为什么实际上大量的实际问题没有用具时滞的模型去描述它们的原因. 但也有不少问题像 (2.1.9) 中描述的情形一样，时滞 T 相当大，是不可以忽略的，必须用具时滞的模型去描述它们.

4. 考虑年龄因素的人口模型

假定在一个稳定社会中研究人口发展过程. 设人口的数量不仅和时间 t 有关，而且还和年龄有关. 如果人口总数很大，那么就可以将人口数量的分布密度视为随时间 t 及年龄 x 连续演化的状态变量 $p(t,x)$, 下面建立它满足的连续时间演化方程.

先看一下人口数量的分布密度 $p(t,x)$ 与时刻 t 时的人口总数 $P(t)$ 的相依关系. 注意在时刻 t, 年龄在 $[x, x+\mathrm{d}x]$ 中的人口数为 $p(t,x)\mathrm{d}x$. 因此在时刻 t 时的人口总数为

$$P(t) = \int_0^\infty p(t,x)\mathrm{d}x.$$

实际上，上式右端中的积分上限只要取人的最大年龄.

现在考虑一个特殊情形, 即不考虑死亡的情形. 那么对任何人来说，在时刻 $t+\mathrm{d}t$, 年龄在 $[x, x+\mathrm{d}x]$ 中的人数应当等于在时刻 t, 年龄在 $[x-\mathrm{d}t, x-\mathrm{d}t+\mathrm{d}x]$ 中的人数，而前者为 $p(t+\mathrm{d}t, x)\mathrm{d}x$, 后者为 $p(t, x-\mathrm{d}t)\mathrm{d}x$, 从而有

$$p(t+\mathrm{d}t, x) = p(t, x-\mathrm{d}t).$$

上式可化成如下形式:

$$[p(t+\mathrm{d}t, x) - p(t,x)] - [p(t, x-\mathrm{d}t) - p(t,x)] = 0,$$

从而易知 $p(t,x)$ 满足方程

$$\frac{\partial p}{\partial t}(t,x) + \frac{\partial p}{\partial x}(t,x) = 0. \tag{2.1.10}$$

但是实际上必须考虑死亡的影响，假设与年龄有关的死亡率 $d(x)$ 与时间 t 无关，那么可得 $p(t,x)$ 满足方程

$$\frac{\partial p}{\partial t}(t,x) + \frac{\partial p}{\partial x}(t,x) = -d(x)p(t,x). \tag{2.1.11}$$

(2.1.10) 与 (2.1.11) 均是一阶线性偏微分方程, 即它们是时空动力学模型.

由这个例子可见, 在考虑不同因素时, 将得到不同的动力学模型, 而具体地看:

(1) 状态变量随时间演化的规律是离散的时刻还是连续的时间;

(2) 状态变量仅依赖于时间演化还是另外依赖于其它的变量;

(3) 状态变量在 t 时刻的变化率仅与 t 时刻的状态变量的值有关, 还是与状态变量在 t 时刻以前一段时间的值有关.

前面介绍的均是定常的 (或称自治的) 系统, 实际上还有许多情形归结的数学模型是非定常 (或称非自治的). 下面章节中将有这方面的内容的介绍.

2.1.2 离散动力学模型与问题

下面先研究一维离散动力学模型, 然后对二维以及高维的离散动力学模型进行研究.

1. 一维离散动力学模型与问题

一维离散动力学是描述如下的函数迭代的离散过程

$$x, f(x), f^2(x), \cdots, f^n(x), \cdots.$$

函数这个概念几何上可以理解为映射, 它是把一个点 (自变量) 对应于另一个点 (因变量) 的几何过程. 由函数 $f(x)$ 确定的一维动力系统也可看成映射 $f(x)$ 确定的一维动力系统, 这样把离散的动力学演化过程看成了点的状态不断映射的变化过程. 函数迭代过程也可以写成差分方程的形式.

(1) 周期点及其吸引排斥性

以下部分介绍离散动力系统的一些基本概念: 前向轨道、后向轨道、全轨道、不动点、周期点、周期点的稳定集、周期点的不稳定集等.

定义 2.1.1 点 $x, f(x), f^2(x), \cdots$ 的集合称为 x 的 **前向轨道**, 用 $O^+(x)$ 表示. 如果 f 是同胚, 即 f 是一对一的、连续的、逆映

射 $f^{-1}(x)$ 也连续, 那么 $x, f^{-1}(x), f^{-2}(x), \cdots$ 称为 x 的 **后向轨道**, 用 $O^-(x)$ 表示. $f^n(x)$ ($n \in \mathbf{Z}$) 的点的集合称为 x 的 **全轨道**, 用 $O(x)$ 表示.

我们研究离散动力系统的基本目的是为了知道映射的全部轨道, 特别是轨道变化的终态. 通常全轨道和前向轨道可能是十分复杂的集合, 大致可以分为周期轨道与非周期轨道两类.

定义 2.1.2 f 的 **不动点** x^* 指满足 $f(x^*) = x^*$. 点 x 是映射 f 的 **周期点** 指成立 $f^n(x) = x$. 满足 $f^n(x) = x$ 的最小正数 n 称为 x 的 **真周期**. 用 $\mathrm{Per}_n(f)$ 表示周期为 n (不必是真周期) 的周期点的集合. **不动点集** 即是周期为 1 的周期点的集合, 用 $\mathrm{Fix}(f)$ 表示之. **周期轨道** 是由周期点的所有迭代构成的, 周期为 n 的周期轨道为

$$x, f(x), \cdots, f^{n-1}(x), x, f(x), \cdots, f^{n-1}(x), \cdots.$$

定义 2.1.3 点 x 是 **最终以 n 为周期** 的指存在 $m > 0$, 使得对 $i \geqslant m$, $f^i(x)$ 是以 n 为周期的, 即成立 $f^{n+i}(x) = f^i(x)$, 其中 $i \geqslant m$.

定义 2.1.4 设 p 是以 n 为周期的周期点. 如果 $\lim\limits_{n\to\infty} f^n(x) = p$, 那么称点 x **前向渐近于** p. p 的 **稳定集** 指所有前向渐近于 p 的点构成的集合, 用 $W^s(p)$ 表示.

例 2.1.1 映射 $f(x) = x^2 - 1$ 有两个不动点 $\dfrac{1+\sqrt{5}}{2}, \dfrac{1-\sqrt{5}}{2}$. 而点 0 和 -1 位于周期 2 的周期轨道上.

例 2.1.2 映射 $f(x) = x^3$ 仅有 3 个不动点 $0, 1, -1$, 且 $W^s(0)$ 是开区间 $(-1, 1)$, 且无其它任何的周期点.

例 2.1.3 设 S^1 表示平面上的单位圆, S^1 上的点用对应的极坐标上以弧度为单位来测量的角值 θ 表示. 对任意整数 k, 形如

$\theta+2k\pi$ 的任意角值确定同一个点. 注意到在圆上有 $f(\theta+2\pi)=f(\theta)$, 现定义映射 $f(\theta)=2\theta$, 那么 f 的 n 周期点满足 $f^n(\theta)=2^n\theta=\theta+2k\pi\ (0\leqslant k\leqslant 2^n)$, 即 $\theta=\dfrac{2k\pi}{2^n-1}$. 易知 0 是不动点, 而 $\theta_0=\dfrac{2k\pi}{2^n}$ (n 是自然数, k 是整数) 是最终不动的.

在知道了离散动力系统的不动点或 $n\ (n\geqslant 2)$ 周期点以后, 紧接着要解决的问题是当初始状态不取在该不动点或 n 周期点时, 多次迭代以后会不会越来越接近该不动点或 n 周期点? 用比较专用的名词即不动点或 n 周期点的吸引性与排斥性哪一种情形发生?

定义 2.1.5 设 p 是以 n 为真周期的周期点, 称点 p 是 **双曲的**或 **具有双曲性** 的是指 $|(f^n)'(p)|\neq 1$, 数 $(f^n)'(p)$ 称为周期点的**乘子**.

定义 2.1.6 设 p 是周期为 n 的双曲周期点且 $|(f^n)'(p)|<1$, 则点 p 称为 **吸引周期点** 或 **周期吸引子**; 对于不动点 p 且 $|f'(p)|<1$, 则点 p 称为 **不动点吸引子**. 对于满足 $|(f^n)'(p)|>1$ 的 n 周期点 p, 称为 **周期排斥子**; 对于不动点 p 且 $|f'(p)|>1$, 则点 p 称为 **不动点排斥子**.

例 2.1.4 映射 $f(x)=-\dfrac{1}{2}(x^3+x)$, 易知 0 是唯一的不动点. 由于 $f'(0)=-\dfrac{1}{2}$, 所以 0 是双曲的不动点. 又 1, -1 位于周期为 2 的周期轨道上. 因为 $(f^2)'(\pm 1)=f'(\mp 1)f'(\pm 1)=4$, 所以周期 2 点 1, -1 为双曲的. 当初始状态取在 $(-1,1)$ 中时, 多次迭代后状态趋向于 0 点, 即 0 不动点是吸引的. 对于 2 周期点 1, -1, 当实轴上的其它点作为初始状态时, 多次迭代以后是远离周期点 1, -1 的, 即它们是排斥的.

例 2.1.5 映射 $f(x)=\dfrac{1}{2}(x^3+x)$, 易知 0, 1, -1 是三个不动点, 因为 $f'(0)=\dfrac{1}{2}$, $f'(\pm 1)=2$, 所以这些不动点均是双曲的. 当初始状态取在 $(-1,1)$ 中时, 多次迭代后趋向于 0, 即不动点 0 是

吸引的，也就是 $(-1,1)$ 是不动点 0 的稳定集 $W^s(0)$. 但是对整个实轴 R 上除 ±1 以外的点作为初始状态，迭代的结果是离开不动点 ±1, 而到 $\pm\infty$ 或到 0 点，即不动点 ±1 是排斥的.

(2) 非周期点及奇异吸引子

考虑本节介绍的虫口模型 (2.1.3), 当参数 $\lambda \in [\lambda_\infty, 4]$ 时，其中 $\lambda_\infty = 3.569945673\cdots$, 此时，绝大部分情形下，状态变量 x_n 的渐近状态是非周期的. 以 $\lambda = 4$ 为例，即 (2.1.3) 为

$$x_{n+1} = 4x_n(1 - x_n). \tag{2.1.12}$$

首先看一下计算机数值计算的结果. 如果初始 x_0 取为 0.1 或 0.100 000 01 或 0.100 000 02 三种非常接近的值，这种接近程度对一般的计算机是不可避免的误差造成的，那么按 $y = 4x(1-x)$ 的迭代函数计算 $x_1, x_2, \cdots$, 一般人们凭想象会认为一直会得到很接近的 3 组值. 让我们看实际迭代吧！奇怪的是在迭代次数小于 50 时，想象是对的，而迭代次数达到 50 及比 50 大时，就会看到原来想象是不对的，见表 2.1.1.

表 2.1.1　$y = 4x(1-x)$ 迭代对初值极端敏感依赖性

n	第一组 x_n	第二组 x_n	第三组 x_n
0	0.1	0.100 000 01	0.100 000 02
1	0.36	0.360 000 003 2	0.360 000 006 4
2	0.921 6	0.921 600 055 8	0.921 600 071 1
$\cdots$	$\cdots$	$\cdots$	$\cdots$
10	0.147 836 559 9	0.147 824 444 9	0.147 812 518 2
$\cdots$	$\cdots$	$\cdots$	$\cdots$
50	0.277 569 081 0	0.435 057 399 7	0.055 005 377 6
51	0.802 094 386 2	0.983 129 834 6	0.207 919 144 2
52	0.634 955 927 4	0.066 342 251 5	0.658 755 094 6

从表中可见，x_{50} 的值 3 组相差很大，又很难预测 x_n $(n>50)$ 的值．看上去 $x_{50},x_{51},x_{52},\cdots$ 即第 50 次迭代及以后迭代的值构成的轨道是非周期的，且像是随机的，其大小又预测不到，给一个名称为 **混沌轨道** 或 **奇异吸引子**.

我们从这个实例可见奇异吸引子有以下 3 个特征：第一个特征是对初值的极端敏感性，初值的小的偏差扰动会引起以后的大的变化或改变．第二个特征是长期不可预测的，虽然从 x_{50} 的 3 组值预测 x_{51} 的 3 组值即短期可以预测的，但是若问 x_{100} 的 3 组值将会怎样，则难以回答了．第三个特征虽然模型是确定的，但是状态的动态变化值好像是随机的.

下面让我们分析一下为什么会造成混沌的，特别是对初值的极端敏感性是怎么产生的，从信息的角度如何理解它们？

现在用下面的例子来分析．考虑

$$f(x)=\begin{cases}2x, & \text{如果 } x\in\left[0,\dfrac{1}{2}\right),\\ 2x-1, & \text{如果 } x\in\left[\dfrac{1}{2},1\right),\end{cases}\tag{2.1.13}$$

迭代映射 $x_{n+1}=f(x_n)$，取初值 $x_0\in[0,1)$，如果把 $[0,1)$ 中的数 x 用二进制小数表示，那么映射 (2.1.13) 就是把 x_n 的二进制小数表示的数中的小数点向右挪一位，而后取其小数部分，即成 x_{n+1}.

例如十进制数 $x_0=\dfrac{11}{32}$，写成二进制小数形式为 $x_0=0.01011$，那么经过 f 的 1 次至 5 次映射依次得到 $x_1=0.1011$, $x_2=0.011$, $x_3=0.11$, $x_4=0.1$, $x_5=0$．易知 $x_n=0$ $(n\geqslant 5)$，即迭代的解趋向于不动点 0.

又如十进制数 x_0 取为 $\dfrac{1}{\sqrt{2}}$，相应的二进制小数表示为

$$x_0=0.1011010000010011\cdots.$$

而 x_5 大于 $\dfrac{1}{2}$ 还是小于 $\dfrac{1}{2}$ 取决于 x_0 中第 6 位小数，现在 x_0 中第 6 位小数是 1，故 $x_5>\dfrac{1}{2}$，又 $x_{12}<\dfrac{1}{2}$ 因为 x_0 中第 13 位小数是 0．一般地，x_k 是大于 $\dfrac{1}{2}$ 还是小于 $\dfrac{1}{2}$ 取决于 x_0 中的第 $k+1$ 位

小数是 1 还是 0. 要知道 x_n 以后的各个值的运动情况，则必须知道 x_0 中的第 $n+1$ 项以后各位小数，而前 n 位小数 (不管是 1 还是 0) 在运算过程中都被一笔抹煞了. 回到现实的数值计算，如果是用计算机来进行数值运算，初值只可能精确到有限位小数，例如设想已精确到 3000 位二进位约合 900 位十进位，那么经过有限次运算 (约 3000 次后)，初值的原始信息在迭代中已消耗殆尽. 如果这时候计算机还照样输出什么结果，这个结果就和 x_0 毫无关系，自然不代表由 x_0 确定的真实运动，而只是由于一些偶然因素 (由计算机构造确定) 而输入的信息引起的. 数学上说，$x_0=$ 无理数时，$n>3000$ 时 x_n 虽不循环，总体看来杂乱无章，但实际上还是确定的，它类似于用计算机制造的 "伪随机数". 但是既然计算机无法分辨差别小于机器精度的两个初值，得到的结果已不是伪随机数，而是真正的随机数. 由于计算精度有限，对初值依赖的极端敏感性意味着内在的随机性.

对于迭代映射 (2.1.13)，当相邻的两个初值 $x_0^{(1)}, x_0^{(2)}$ 落在 $\left[0,\frac{1}{2}\right]$ 或 $\left[\frac{1}{2},1\right]$ 的两侧时，得到的相应的两个迭代值 $x_1^{(1)}$, $x_2^{(2)}$. 一种情形是间隔距离拉长 1 倍，但仍落在初值所在同一侧，即 $\Delta x_1 = x_1^{(2)}-x_1^{(1)}=2\Delta x_0=2(x_0^{(2)}-x_0^{(1)})$. 另一种情形，迭代后的值可能落在初值所在的另一侧，或跨两侧，一侧一个迭代点，此时可见以后迭代下去点的趋向就不再是简单地扩大两倍，也就造成像混沌这样的复杂动态. 从局部来说，每一处的特征值 (即斜率) 都是 2，误差愈放愈大 (普遍意义下的不稳定)；而从整体来说，全局的运动又不准离开单位正方形. 形象地说，局部不稳定而全局是有界的、有限制的，于是既 "跑不掉" 又 "稳不住" 或 "无处安家"，只能在被限制的范围内到处游荡着. 现从另一个角度来理解，即一方面拉伸，另一方面又折叠，就造成了混沌现象. 将迭代映射 (2.1.13) 用下面称为相图的图 2.1.1 来形象描述，图中两个直线段，每段的斜率均是 2. 现取初始值及扰动的初始值之间隔为 Δx_0. 在左侧，迭代一次得 Δx_1，仍在同一侧，但放大了 1 倍. 再迭代一次得 Δx_2，

此时迭代的值跨两侧, 即不在同一侧. 接下去迭代得 Δx_3, 它不再是 $2\Delta x_2$, 即把误差扩大 2 倍已不成立.

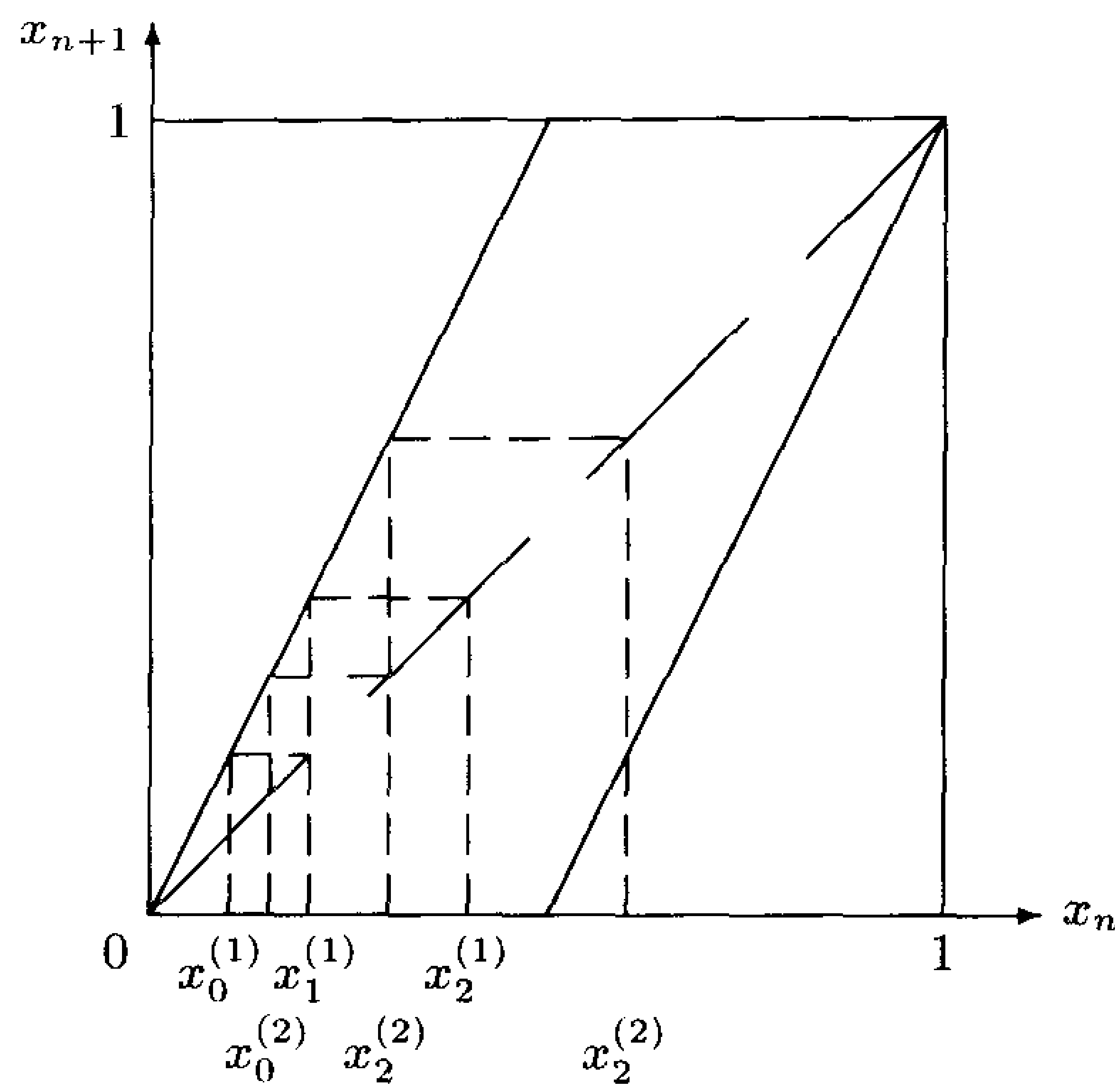


图 2.1.1　导致混沌的迭代 (2.1.13) 的相图

2. 高维的离散动力学模型与问题

首先介绍一个二维离散动力学模型的例子 —— Henon 映射, 其次考虑一般的 n 维离散动力系统的模型以及一般概念, 最后介绍这些模型研究中的问题.

(1) 一个二维的例子 —— Henon 映射

法国天文学家 Henon 为了模拟气象学家 Lorenz 研究的著名方程出现的复杂吸引子, 抛开了连续的时间演化的微分方程而改用离散时间演化的二维映射

$$\begin{cases} x_{n+1} = 1 + y_n - a x_n^2, \\ y_{n+1} = b x_n, \end{cases} \tag{2.1.14}$$

其中 a, b 为参数, 这里仅 x_n^2 为非线性项.

现在讨论参数值 $a = 1.4$, $b = 0.3$ 的情形.

类似于考虑线性映射的不动点稳定性可以应用特征值分析, 对非线性映射应用 Jacobi 行列式及矩阵特征值的分析是很有用的

方法. 注意 (2.1.14) 的不动点有两个:

$$A(0.631, 0.631), \quad B(-1.131, -1.131),$$

实际上不动点 (ξ, η) 满足方程组

$$\begin{cases} \xi = 1 + b\eta - a\xi^2, \\ \eta = b\xi, \end{cases}$$

而不动点的稳定性与不稳定性是由映射的变分矩阵在该不动点处的两个特征值的模小于 1 与大于 1 来判别的. 如果两个特征值的模全部小于 1, 则不动点是稳定的, 即吸引的; 如果至少有一个特征值的模大于 1, 则不动点是不稳定的, 即排斥的. 对现在讨论的情形, 在 A 点和 B 点处变分矩阵的特征值的模一个大于 1, 一个小于 1, 即为鞍点, 两个不动点均为不稳定的、排斥的. 此时 Jacobi 行列式

$$J = \left| \frac{\partial(x_{n+1}, y_{n+1})}{\partial(x_n, y_n)} \right| = \begin{vmatrix} -2ax_n & 1 \\ b & 0 \end{vmatrix} = -b.$$

可见当 $|b| < 1$ 时, 有 $|J| < 1$, 即该二维迭代映射是把原有的面积压缩, 称它为 **耗散系统**. 现在讨论的情形属于这类系统. 当然, 若取 $b = -1$, $J = 1$, 则此时的映射变为保面积的, 又称之为 **保积系统**.

对现在讨论的情形, 存在一个奇异吸引子, 它有如下一些性质:

1) 取不同初值, 映射的相平面上轨道整个演化过程均形成大范围几何性能相同的奇异吸引子.

2) 大范围的几何性能. 先取一个称为捕捉区 D 的四边形 $PQRS$, 例如取

$$P(-1.325, 1.39), \quad Q(1.32, 0.45),$$

$$R(1.25, -0.41), \quad S(-1.05, -1.56).$$

下面考虑四边形的四条边 PQ, QR, RS, SP 在 f 映射下变到什么地方去了，从而可发现 $f(D) = P'Q'R'S' \subset D$. 为了弄清楚整个一条边 PQ 是如何变到 $P'Q'$ 的，可以在 PQ 中设置一个点 W, 考察 W 是如何经 f 映射到 W' 的. 考虑 $f(D)$ 可以将 D 的变化分成 3 步，第一步拉长或压扁，使面积缩小为原来的 0.3 倍，这是因为 $|J| = 0.3$. 第二步将之翻身，即 P 与 Q, S 与 R 对接，这是因为 J 为负的. 第三步将之变成弓形，这是因为有了 x^2 项的作用，直线段变成了抛物线段. 因为 D 选取得比较理想，正好 $f(D) \subset D$ (见图 2.1.2).

我们继续考虑 $f^2(D)$. 此时 $P'Q'R'S'$ 已变成在 $f(D)$ 弓形区域中的更细小的弓形 $f^2(D)$. 继续此过程有

$$D \supset f(D) \supset f^2(D) \supset \cdots \supset f^n(D) \supset f^{n+1}(D) \supset \cdots,$$

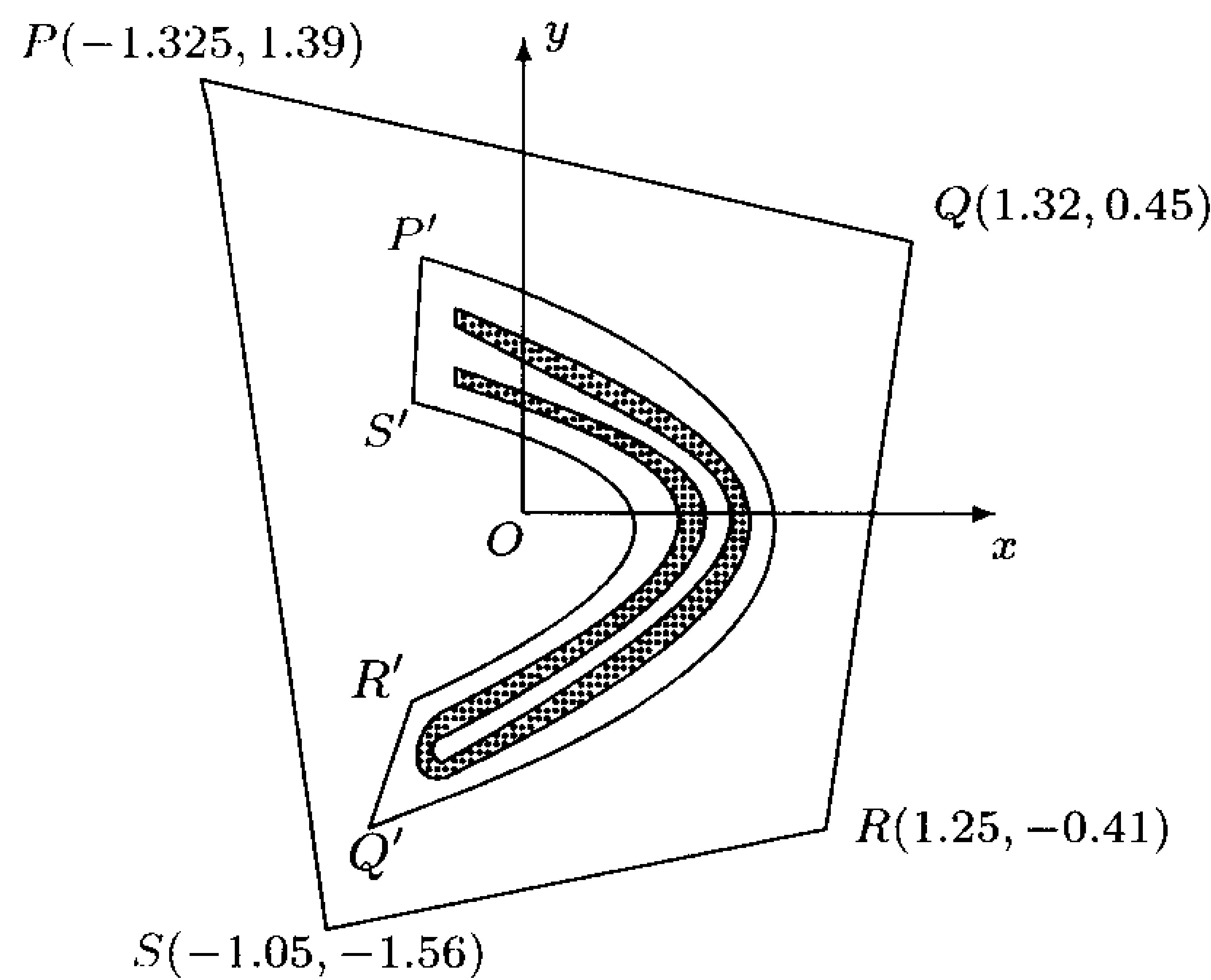


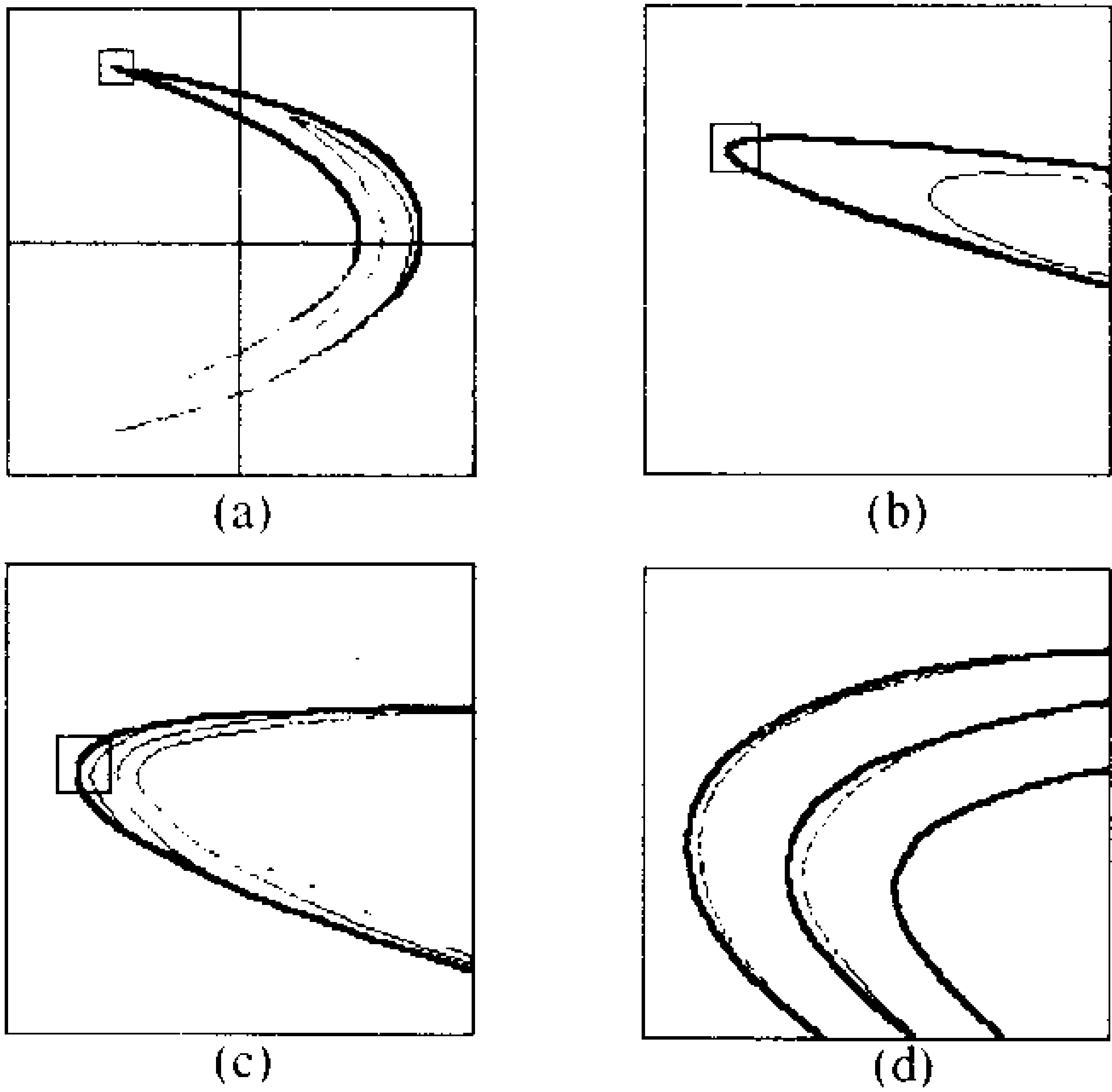
图 2.1.2　捕捉区 D 在映射 $f, f^2, \cdots$ 下的几何特征示意图

最后形成的集合有下列特征：$f^n(D)$ 的面积随 $n \to +\infty$ 而趋于 0; $\lim\limits_{n\to\infty} f^n(D)$ 是一个区域，在 f 的映射下是不变的；这极限区域是一个奇异吸引子，它是一条无限长的无穷多次盘旋迂回的曲线，一个局部进行放大后，其内部结构又有无穷多层次，吸引子的每一个

横截面是一个非正规的 Cantor 集，该吸引子具有无穷多层次的自相似结构，如图 2.1.3 所示，各图小方块内的几何图形放大以后，在下一个图中表示，由图可见其结构是相似的.

注意图 2.1.3(a) 中的小方块放大即为图 2.1.3(b), 依图 (b) 的小方块放大可见图 2.1.3(c), 图 (c) 的小方块放大可见图 2.1.3(d). 这些放大图很好地说明了 Henon 吸引子的分形特性，即奇异吸引子在所有尺寸都具有类似的结构. 事实上，奇异吸引子本身就是分形集.

3) 固定 $b=0.3$, 仅变参数 a. 当 a 增大时，出现周期解，并有倍周期化. 在 $a_1=0.3675$ 时，周期 1 分叉为周期 2; $a_2=1.0$, 周期 2 分叉为周期 4; $a_3=1.01$, 周期 4 分叉为周期 8, $\cdots$; 到 $a_\infty=1.058$ 时，进入混沌区； $a=1.153$ 时，两片混沌变成一片；到 $a>1.427$ 时，不变集合不再是吸引的了.



(a) (b) (c) (d)

图 2.1.3　Henon 映射的奇异吸引子

当然以上分析是数值模拟计算的结果，是定性分析的结果. 奇异吸引子的定义是描述性的、不严格的，应用时要注意尽量用 §2.2 中介绍的数量分析的方法去验证这些定性结果.

(2) 一般的高维离散动力学模型

现在用如下的一般的数学形式来表示 n 维的离散动力学模型

$$\boldsymbol{x}(t+1)=\boldsymbol{f}(\boldsymbol{\alpha},\boldsymbol{x}(t)),\quad (t=1,2,\cdots,n), \tag{2.1.15}$$

其中 $\boldsymbol{x}(t)\in \mathrm{R}^n$, $\boldsymbol{\alpha}\in \mathrm{R}^m$, $\boldsymbol{f}:\mathrm{R}^m\times \mathrm{R}^n\to \mathrm{R}^n$, $\boldsymbol{x}(t)$ 是状态变量向量, $\boldsymbol{\alpha}$ 是参数向量. 这个模型是定常的或称为自治的. 如果 $\boldsymbol{f}$ 还依赖于 t 或者参数 $\boldsymbol{\alpha}$ 依赖于 t, 那么模型是不定常的或称为非自治的. 我们约定以后 (2.1.15) 中向量记号略去.

对模型 (2.1.15) 从另外一个角度来看, 建立初值 $x(0)$ 及迭代次数 n 与第 n 步迭代值 $x(n)$ 之间的映射 $\pi(n,x(0))$, 即改记 (2.1.15) 为

$$x(n)=\pi(n,x(0)). \tag{2.1.16}$$

数学上 R^n 上的离散动力系统的定义为映射 $\pi:\mathrm{Z}^+\times\mathrm{R}^n\to\mathrm{R}^n$, 若它对一切 $m,k\in\mathrm{Z}^+$ $(\mathrm{Z}^+=\{0,1,2,\cdots\})$ 及 $x\in\mathrm{R}^n$ 满足下述三个条件:

(i) $\pi(0,x)=x$; $\qquad (2.1.17)_1$

(ii) $\pi(m,\pi(k,x))=\pi(m+k,x)$; $\qquad (2.1.17)_2$

(iii) π 是连续的映射. $\qquad (2.1.17)_3$

事实上, 也可以改记 (2.1.15) 为

$$x(m)=T^m(x(0)), \tag{2.1.18}$$

这里 T^m 是 T 的 m 重迭代, 而 T 为 $f(\alpha,\cdot)$, $T(x)=f(\alpha,x)$, 即有 $T(x)=\pi(1,x)$.

这些从不同角度来表示 (2.1.15) 的形式在研究动力学问题时有它们的特殊意义.

2.1.3 连续动力学模型与问题

对于用常微分方程描述的动力学模型, 有自治的与非自治的之分, 即定常的与非定常的之分.

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(\alpha,x), \tag{2.1.19}$$

其中 $x \in \mathrm{R}^n, \alpha \in \mathrm{R}^m, f: \mathrm{R}^m \times \mathrm{R}^n \to \mathrm{R}^n$ 是连续的，这是一个 **定常的** 或称为 **自治的连续的动力学模型**, $\alpha \in \mathrm{R}^m$ 称为 **参数**, $x \in \mathrm{R}^n$ 称为 **状态变量**.

(2.1.19) 的初值问题为：求满足下述初始条件

$$x(t_0) = x_0 \in \mathrm{R}^n \tag{2.1.20}$$

的 (2.1.19) 的解. 如果 (2.1.19), (2.1.20) 的解 $\varphi(t, x_0)$ 的存在区间是 $(-\infty, +\infty)$, 即 $\varphi(t, x_0)$ 对一切 $t \in \mathrm{R}$ 和 $x_0 \in \mathrm{R}^n$ 有定义，现把 x_0 改写为 x, 那么满足下面两个条件的解映射 $\varphi: \mathrm{R} \times \mathrm{R}^n \to \mathrm{R}^n$ 称为 R^n 中的 **连续动力系统** 或 **流**:

(i) $\varphi(0, x) = x, \quad \forall x \in \mathrm{R}^n;$ $\qquad (2.1.21)_1$

(ii) $\varphi(s + t, x) = \varphi(s, \varphi(t, x)), \quad \forall s, t \in \mathrm{R}, x \in \mathrm{R}^n.$ $\qquad (2.1.21)_2$

对于给定的 $x \in \mathrm{R}^n$, 称点集

$$\mathrm{Orb}_{\varphi}(x) = \{\varphi(t, x) \mid t \in \mathrm{R}\} \subset \mathrm{R}^n \tag{2.1.22}$$

为流 φ 经过点 x 的 **轨道**.

1. 相空间与相平面

为了讨论 (2.1.19), (2.1.20) 的解的性质或形如 (2.1.22) 的轨道的性质，引进相空间与相平面的概念，它们提供了在实际上很重要的想法与工具.

如果要在几何上表示解 $\varphi(t, x_0)$，首先想到的是把解轨线在 (t, x) 的 $n+1$ 维空间中进行几何描述. 但是当 $n > 1$ 时，就得至少在三维空间中描述之. 这是比较复杂的甚至是不可能的. 现退一步把状态变量 x 所在的空间 R^n 作为几何描述轨线的相空间，如果 $n = 2$, 即二维系统，则 R^2 为平面情形，即相空间实际上为相平面. 此时几何地表示轨线比较简单且形象直观. 但对 $n \geqslant 3$, 则同样存在复杂性，甚至很难讨论，此时可以仍用 n 个状态变量中的某两个状态变量，例如 x_i, x_j $(i, j = 1, 2, \cdots, n)$ 构成相平面 Ox_ix_j, 在这个相平面中显示以参数 t 表示的曲线

$$x_i = x_i(t), \quad x_j = x_j(t).$$

这样便可以简单地实现将轨线 $x = \varphi(t, x_0)$ 在相平面上的投影表示出来. 易见二维相平面是典型的情形.

2. 平衡位置及其稳定性

考虑简单的二维线性自治系统

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = ax + by, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = cx + dy. \end{cases} \tag{2.1.23}$$

点 $(0,0)$ 称为 (2.1.23) 的 **平衡位置**, 又称为 **平衡点**. 置 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, 一般设 $\det A \neq 0$. 由 A 的特征方程

$$\lambda^2 - (a+d)\lambda + (ad - bc) = 0$$

得到它的两个特征根 λ_1 和 λ_2, 有时简记为 $\lambda_{1,2}$. 根据特征值的不同情形有不同类型的相平面中的轨线.

由平衡点满足的方程得, 在平衡点处状态变量随时间的变化率为零, 但是轨线并不意味着一定渐近地趋向于平衡点, 或在平衡点处终结, 相反地有时可能是越来越远离平衡点或始终与平衡点保持一个距离. 就相平面的轨线随时间增大的变化趋势而言, 如果所有的轨线均沿着时间增加方向越来越趋向于平衡点, 那么称该平衡点是 **渐近稳定** 的; 如果所有的轨线沿着时间增加方向既不偏离平衡点越来越远, 也不越来越接近平衡点, 即可以与平衡点是接近但不能趋向于平衡点, 那么称该平衡点是 **稳定但不渐近稳定**, 当然渐近稳定的平衡点必稳定; 如果至少存在一条轨线是越来越远离平衡点的, 那么称该平衡点为 **不稳定的**.

二维自治系统的平面轨线图有如下几种:

(1) 渐近稳定的情形

如果两个特征根 λ_1, λ_2 的实部均为负的, 那么按照 $\lambda_{1,2}$ 是共轭复根还是两实根可区分为平衡点是稳定的焦点还是稳定的结点 (见图 2.1.4 和图 2.1.5).

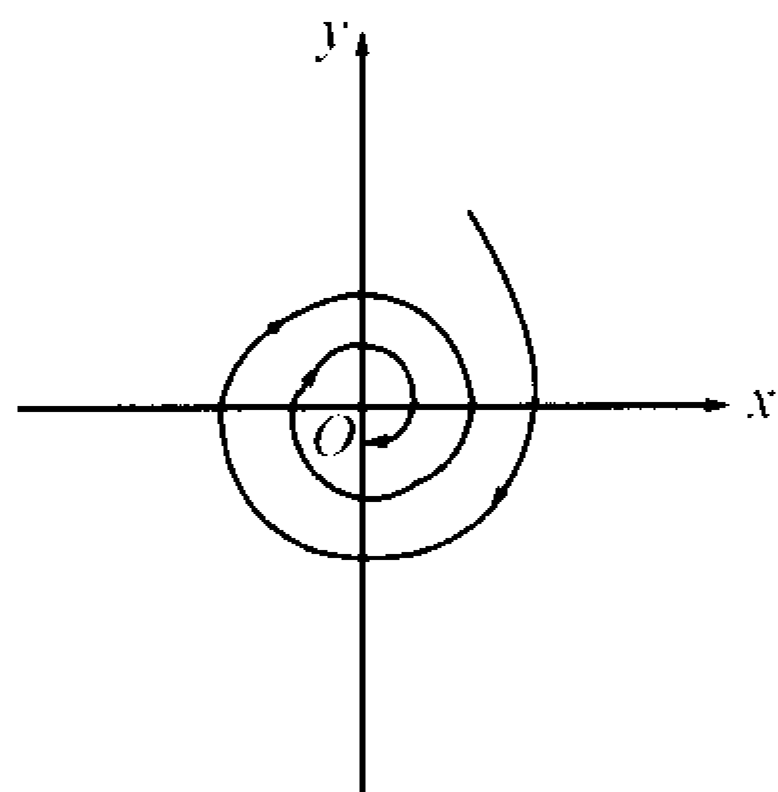


图 2.1.4　稳定的焦点

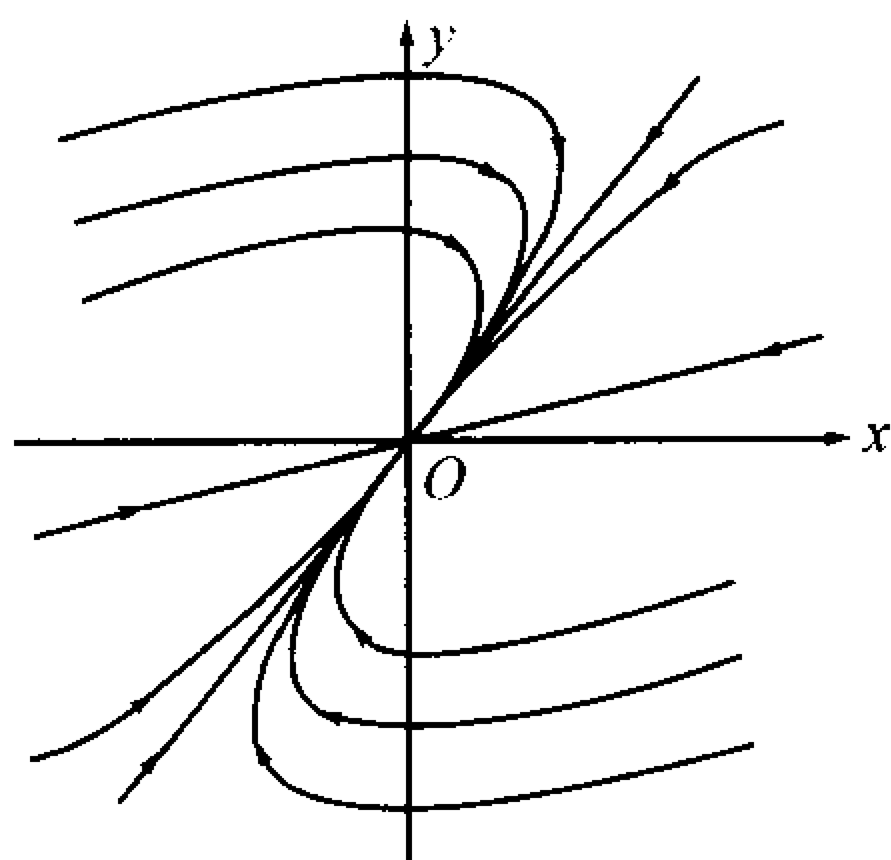


图 2.1.5　稳定的结点

(2) 稳定但不渐近稳定的情形

如果两个特征根是一对纯虚根，那么平衡点仅是稳定的，而不是渐近稳定的，称之为 **中心** (见图 2.1.6).

(3) 不稳定的情形

如果两特征根是一对具有正实部的共轭复根，那么平衡点是不稳定的焦点 (见图 2.1.7); 如果两特征根是两个正实根，那么平衡点是不稳定的结点 (见图 2.1.8); 如果两个特征根一个是正实根，另一个是负实根，那么平衡点是鞍点 (见图 2.1.9).

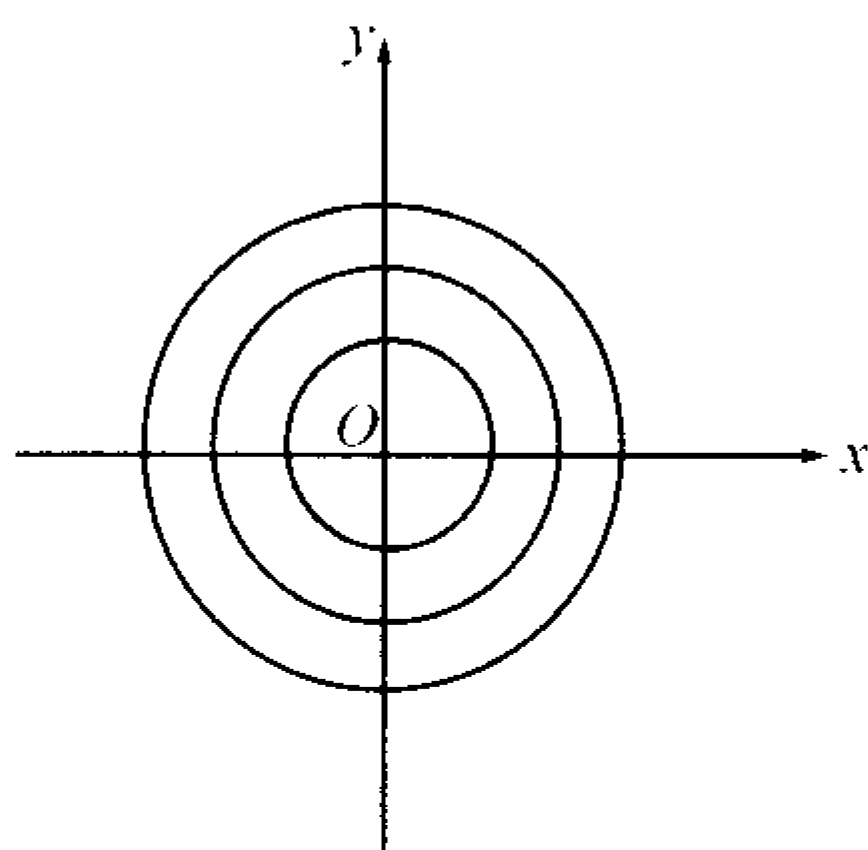


图 2.1.6　中心

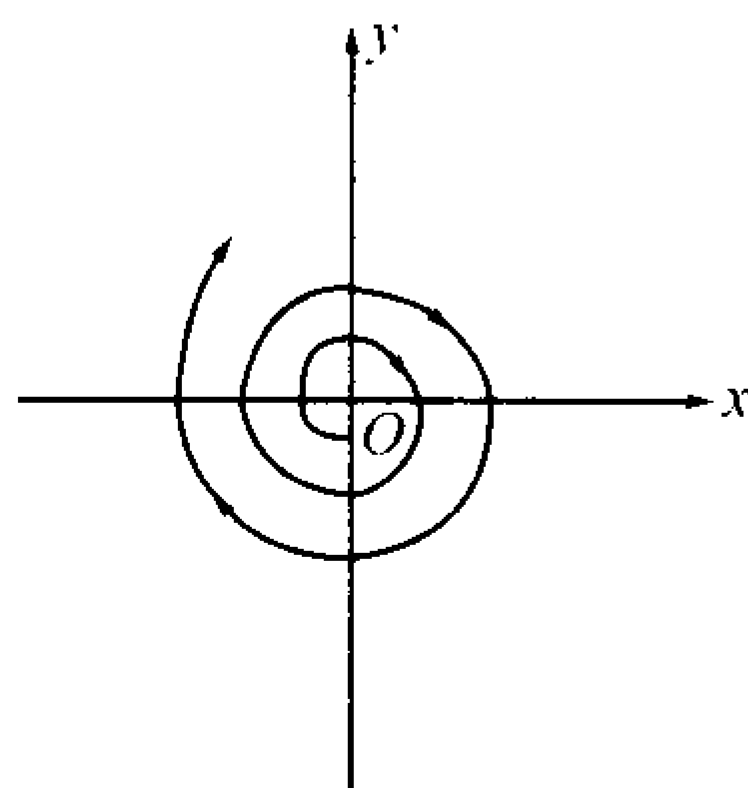


图 2.1.7　不稳定焦点

对于二维的非线性自治系统

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x, y), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = g(x, y), \end{cases} \tag{2.1.24}$$

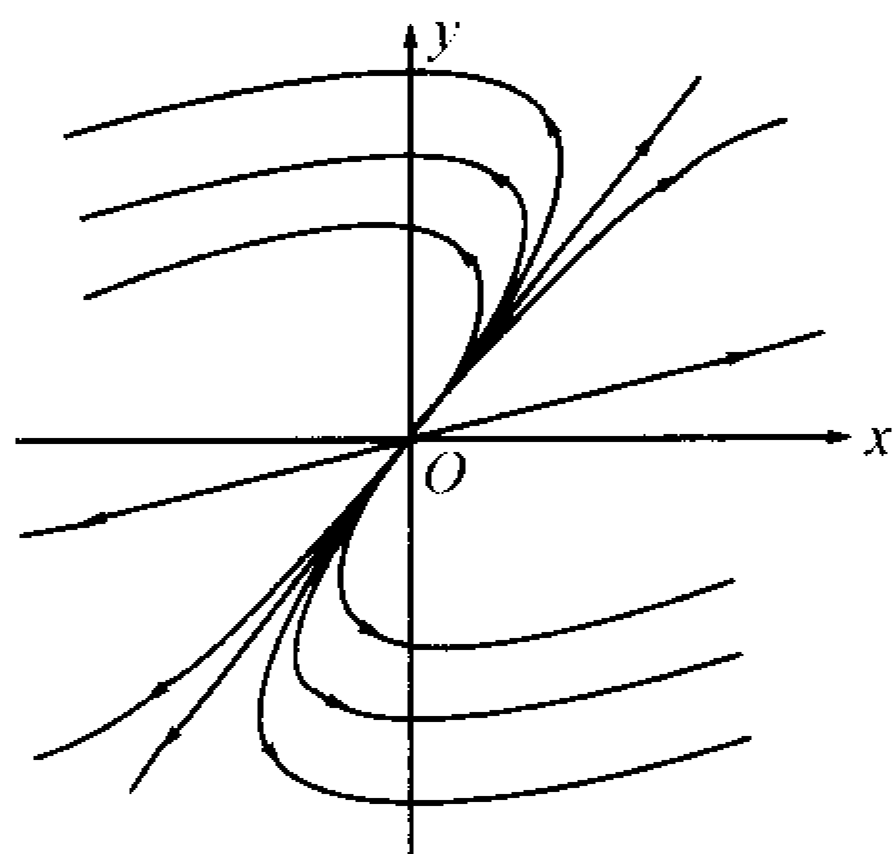


图 2.1.8　不稳定结点

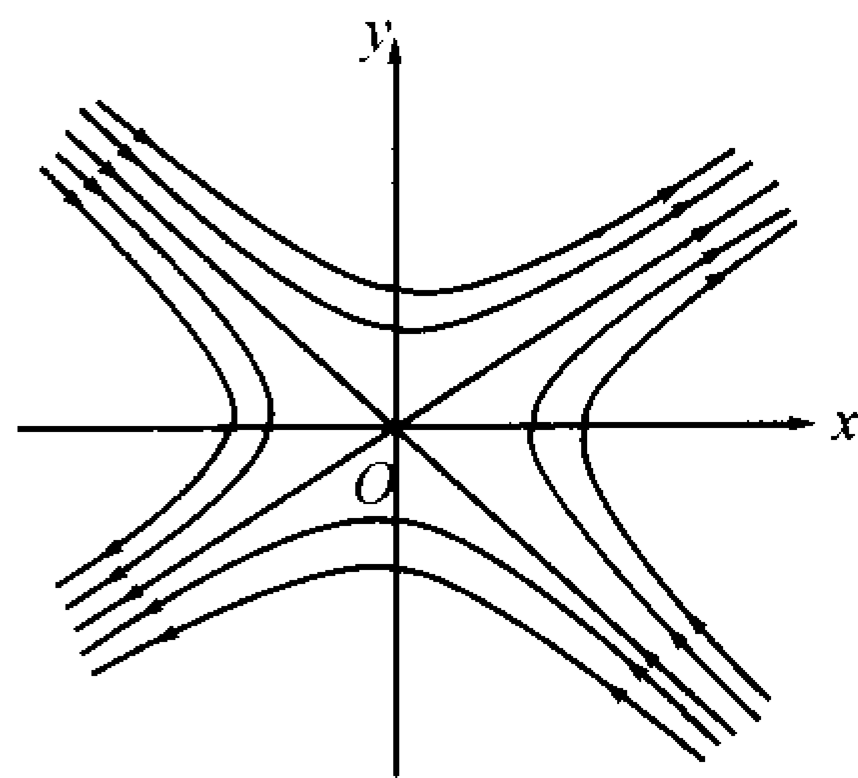


图 2.1.9　鞍点

可以先将其右端函数 $f(x,y)$ 与 $g(x,y)$ 线性化. 一般地假设 $f(0,0)=0$, $g(0,0)=0$, $f(x,y)=o(r^{\varepsilon+1})$, $g(x,y)=o(r^{\varepsilon+1})$, $\varepsilon>0$, $(r\to 0)$, 其中 $r=\sqrt{x^2+y^2}$. (2.1.24) 的平衡点的稳定性可以通过 (2.1.24) 相应的线性化系统的平衡点的稳定性来判别, 仅有一个例外即平衡点是线性化系统的中心情形, 注意此时相平面上的轨线直线段将变弯曲. 这些结果的严格的数学定理常称为 **Perron 定理**.

实际的非线性系统, 即使是最简单的二维自治系统, 可能平衡点不仅仅是 1 个. 看下面的例子.

例 2.1.6　对于无阻尼单摆的运动方程

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}+\frac{g}{l}\sin x=0, \tag{2.1.25}$$

令 $y=\dfrac{\mathrm{d}x}{\mathrm{d}t}$, 将 (2.1.25) 化为如下二维系统

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=y,\\[2mm] \dfrac{\mathrm{d}y}{\mathrm{d}t}=-\dfrac{g}{l}\sin x.\end{cases} \tag{2.1.26}$$

由联立方程组

$$\begin{cases}y=0,\\[2mm] -\dfrac{g}{l}\sin x=0,\end{cases}$$

求得平衡点为 $x = n\pi$ (n 为整数), $y = 0$. 可见平衡点有无穷多个，对每个平衡点可以讨论其稳定性.

(i) 平衡点 $(2k\pi, 0)$ (k 为整数)

作变换 $u = x - 2k\pi$, 得

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t} = y, \\[2ex] \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{g}{l}\sin(2k\pi + u) = -\dfrac{g}{l}\sin u = -\dfrac{g}{l}u + o(u^2), \end{cases}$$

相应的线性系统为

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t} = y, \\[2ex] \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{g}{l}u. \end{cases}$$

该系统的线性化矩阵 $A = \begin{pmatrix} 0 & 1 \\ -\dfrac{g}{l} & 0 \end{pmatrix}$, 有 1 对纯虚根 $\lambda_{1,2} = \pm\sqrt{\dfrac{g}{l}}\mathrm{i}$. 原方程化为如下形式的相平面方程:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{-\dfrac{g}{l}\sin x}{y},$$

相轨线方程为

$$y^2 = 2\frac{g}{l}\cos x + h. \tag{2.1.27}$$

对应于 $|h| > 2\dfrac{g}{l}$, $|h| = 2\dfrac{g}{l}$ 及 $|h| < 2\dfrac{g}{l}$ 三种情形，可作出相轨线分布图 (见图 2.1.10).

(ii) 平衡点 $((2k+1)\pi, 0)$ (k 为整数)

令 $v = x - (2k+1)\pi$, 得到

$$\begin{cases} \dfrac{\mathrm{d}v}{\mathrm{d}t} = y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{g}{l}\sin[(2k+1)\pi + v] = \dfrac{g}{l}\sin v = \dfrac{g}{l}v + o(v^2), \end{cases}$$

相应的线性系统为

$$\begin{cases} \dfrac{\mathrm{d}v}{\mathrm{d}t} = y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = \dfrac{g}{l}v. \end{cases}$$

该系统的系数矩阵为 $\begin{pmatrix} 0 & 1 \\ \dfrac{g}{l} & 0 \end{pmatrix}$, 特征值为 $\lambda_1 = \sqrt{\dfrac{g}{l}}$, $\lambda_2 = -\sqrt{\dfrac{g}{l}}$. 故平衡点 $(0,0)$ 是线性系统的鞍点, 按照 Perron 定理, 平衡点 $((2k+1)\pi, 0)$ 是原系统 (2.1.25) 的鞍点, 整体的相图见图 2.1.10.

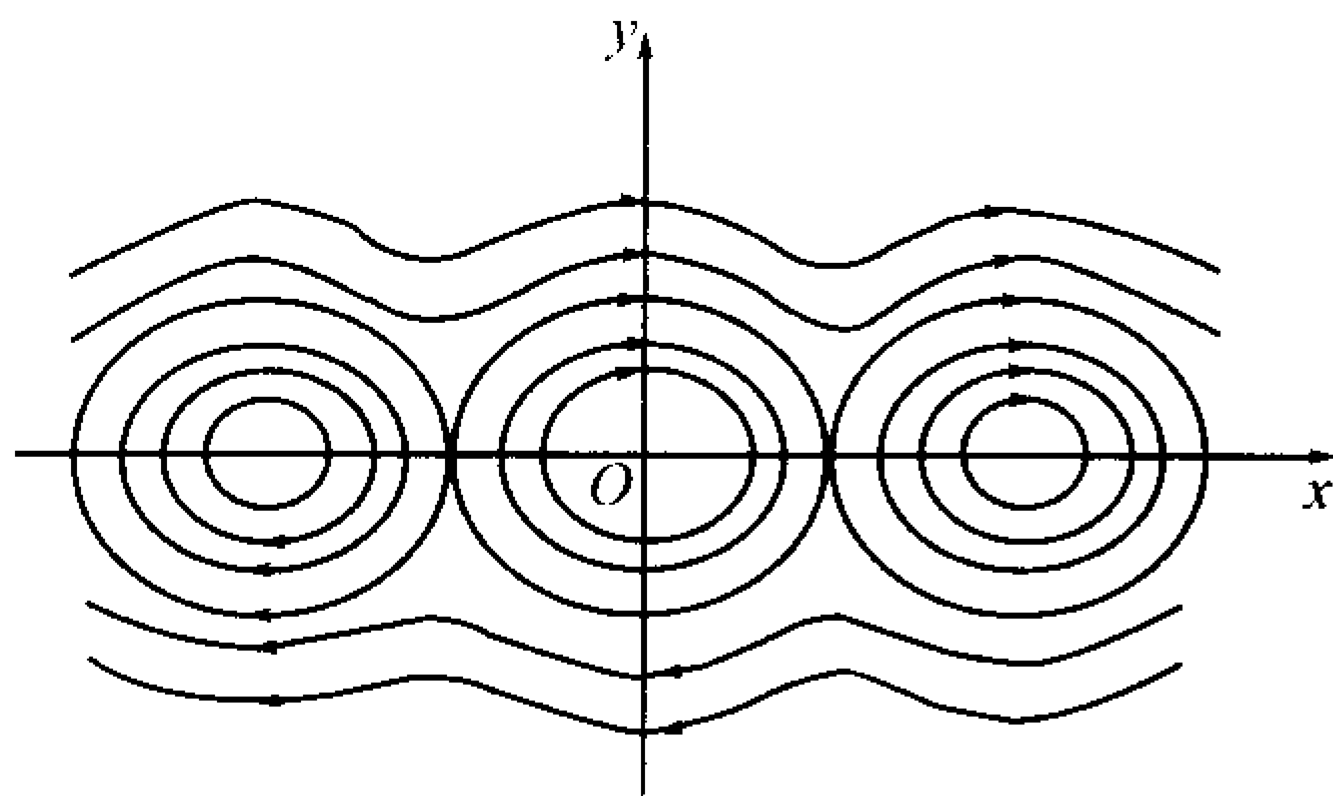


图 2.1.10　系统 (2.1.25) 的相图

对于一般的 n 维系统的平衡点及其稳定性的讨论比较复杂, 将在本章 §2.2 中予以介绍.

3. 极限环 —— 孤立的闭轨

与考虑平衡位置及稳定性一样，先讨论二维的自治系统

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = P(x,y), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = Q(x,y), \end{cases} \tag{2.1.28}$$

考虑它的周期解的几何图像. 设 $x=\varphi(t), y=\psi(t)$ 是 (2.1.28) 的周期为 T 的周期解，即满足

$$\varphi(t+T)=\varphi(t), \quad \psi(t+T)=\psi(t),$$

则该周期解在相平面 $O\text{-}xy$ 平面上的图像 (即相图) 为 1 条闭曲线，称为 (2.1.28) 的 1 条 **闭轨**.

先考虑前面已介绍过的例子

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = -y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = x. \end{cases}$$

相方程为 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{x}{-y}$, 相轨线为 $x^2+y^2=c^2$, 随 $c>0$ 的连续变小趋于零，得到一簇自外向内的包围原点的同心圆. 每个同心圆是一条闭轨，且每条闭轨不是孤立的，它的近旁有许多闭轨存在. 我们下面考虑的闭轨排除这种情形，而关心的是孤立的闭轨. 再看 1 个例子:

例 2.1.7 考虑如下二维非线性自治系统

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = y + x(1-x^2-y^2), \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -x + y(1-x^2-y^2). \end{cases} \tag{2.1.29}$$

引进极坐标 $x = r\cos\theta, y = r\sin\theta$, 利用 $r^2 = x^2 + y^2, \theta = \tan^{-1}\dfrac{y}{x}$, 原系统化为

$$\begin{cases} \dfrac{\mathrm{d}r}{\mathrm{d}t} = r(1-r^2), \\ \dfrac{\mathrm{d}\theta}{\mathrm{d}t} = -1. \end{cases}$$

它的全部解为 $r = 0, r = 1$ 及 $r = \dfrac{1}{\sqrt{1+c\mathrm{e}^{-2t}}}, \theta = -t + \theta_0$, 实际上 $c = 0$ 对应 $r = 1, c \neq 0$ 对应在单位圆内与单位圆外的相轨线, 显见除单位圆外均是螺线形轨线, 即均无其它闭轨线. $c > 0$ 对应位于单位圆内的螺旋线, 当 $t \to +\infty$ 时, $r \to 1$, 即趋向于单位圆周; 对应 $c < 0$ 情形为单位圆周外的轨线, 也是螺旋线的部分段, 当 $t \to +\infty$ 时, $r \to 1$. 由此可知单位圆周是一条孤立闭轨, 数学上称之为 **极限环** 或 **极限圈**, 邻近的轨线当 $t \to +\infty$ 时趋向于单位圆周, 即单位圆周这个极限环是稳定的.

下面给出极限环判别其存在与不存在的环域定理与 Bendixson 和 Dulac 否定孤立闭轨存在的判据.

定理 2.1.1 (环域定理) 设 G 是由两条简单闭曲线 Γ_1 和 Γ_2 围成的环域. 如果在环域 G 中没有方程组 (2.1.28) 的奇点 (单个点的解轨线), 且从 G 的边界上的点出发的轨线都进入 (或离开) G, 则在环域 G 内至少存在方程组 (2.1.28) 的一条闭轨 Γ (见图 2.1.11).

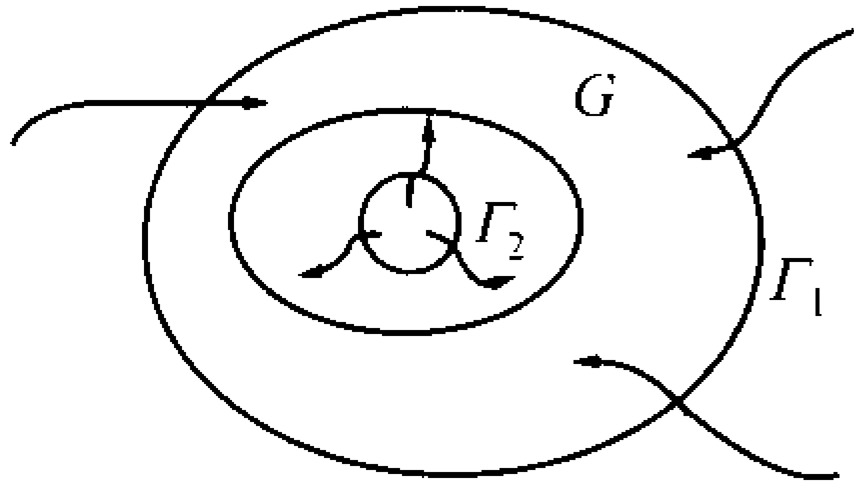


图 2.1.11 环域定理示意图

定理 2.1.2 任何闭轨线所包围的单连通区域内, 至少包含 1

个奇点.

定理 2.1.3 (Bendixson 判据) 设 (2.1.28) 中的函数 $P(x,y)$, $Q(x,y)$ 在单连通区域 G 中存在一阶连续偏导数, 且在 G 中

$$\frac{\partial P(x,y)}{\partial x}+\frac{\partial Q(x,y)}{\partial y}\neq 0, \tag{2.1.30}$$

则微分方程组 (2.1.28) 在区域 G 中不存在闭轨.

定理 2.1.4 (Dulac 判据) 设存在函数 $B(x,y)$, 它和 $P(x,y)$, $Q(x,y)$ 都在单连通区域 G 中有一阶连续的偏导数, 且在 G 中成立

$$\frac{\partial(BP)}{\partial x}+\frac{\partial(BQ)}{\partial y}\neq 0,$$

则微分方程组 (2.1.28) 在区域 G 中不存在闭轨.

n 维 ($n\geqslant 3$) 的自治系统的闭轨及极限环更复杂, 这里暂且不作讨论. 要强调说明的一个问题是对于一个系统, 如果它的解轨线是有界的情形, 则有如下结果:

定理 2.1.5 对二维的有界动力学系统, 在一个有界的单连通区域中, 吸引子的种类只可能是平衡点或极限环.

4. 混沌奇异吸引子

举一个三维非线性自治系统的例子, 说明定理 2.1.5 对 n ($\geqslant 3$) 维的自治系统结论不一定成立, 将可能出现新的吸引子.

例 2.1.8 考虑如下的 Lorenz 方程:

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=-\sigma x+\sigma y,\\[2ex] \dfrac{\mathrm{d}y}{\mathrm{d}t}=-xz+rx-y,\\[2ex] \dfrac{\mathrm{d}z}{\mathrm{d}t}=xy-bz,\end{cases} \tag{2.1.31}$$

此处 σ, r, b 为大于零的参数.

这个方程是由美国麻省理工学院的气象学家 Lorenz 于 1963 年在对流体运动的偏微分方程用 Fourier 级数方法求解时化简归结而得来的. 该系统有 3 个平衡点 (当 $r > 1$ 时):

$$(0,0,0), \quad (\sqrt{b(r-1)}, \sqrt{b(r-1)}, r-1),$$

$$(-\sqrt{b(r-1)}, -\sqrt{b(r-1)}, r-1).$$

原点 (0,0,0) 对应的线性化系统的系数矩阵的特征方程为

$$(\lambda + b)[\lambda^2 + (\sigma+1)\lambda + \sigma(1-r)] = 0,$$

它具有三个实根. 当 $0 < r < 1$ 时, 原点是稳定的平衡点. 但是若取 $b = \frac{8}{3}$, $r = 28$, $\sigma = 10$, 此时 3 个平衡点均为不稳定的. 若初始值选为 (0,1,0), 通过差分格式数值计算得到如下的结果:

(1) 对初值具有敏感依赖性. 将 $y(t)$ 的前 3000 次迭代结果所表示的曲线表示于图 2.1.12 中.

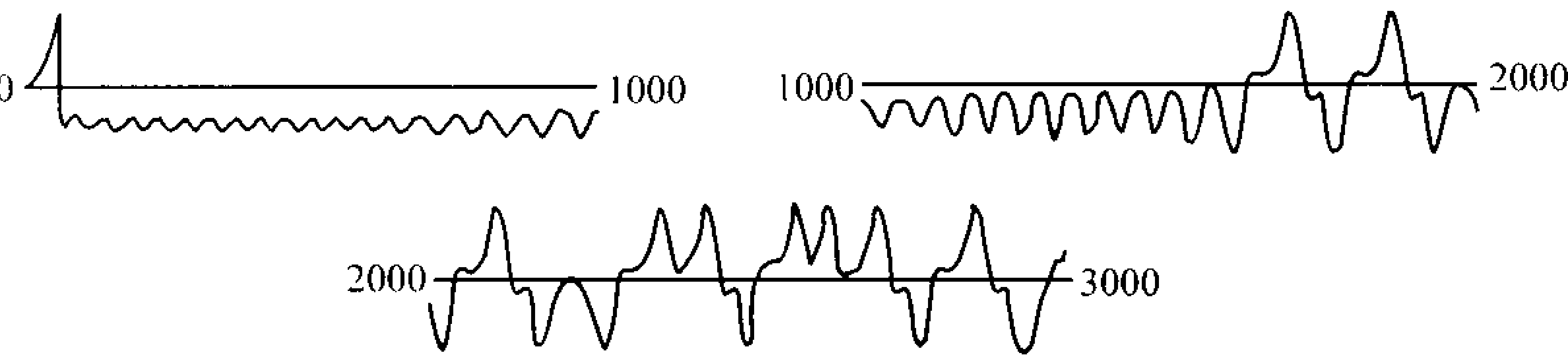


图 2.1.12　$y(t)$ 前 3000 次的迭代结果

(2) 当 $t \to +\infty$ 时, 系统 (2.1.31) 的轨线趋向的最终状态不是平衡点, 也不是极限环, 而是一个测度为零的集合, 是一个奇异吸引子 —— Lorenz 吸引子 (见图 2.1.13).

Lorenz 吸引子由两片构成, 每一片围绕系统的一个平衡点, 系统的轨线在其中的一片上由外向内绕到平衡点附近, 突然跳到另一片的外沿, 由外向内绕, 然后再随机地跳回原来的一片 $\cdots\cdots$.

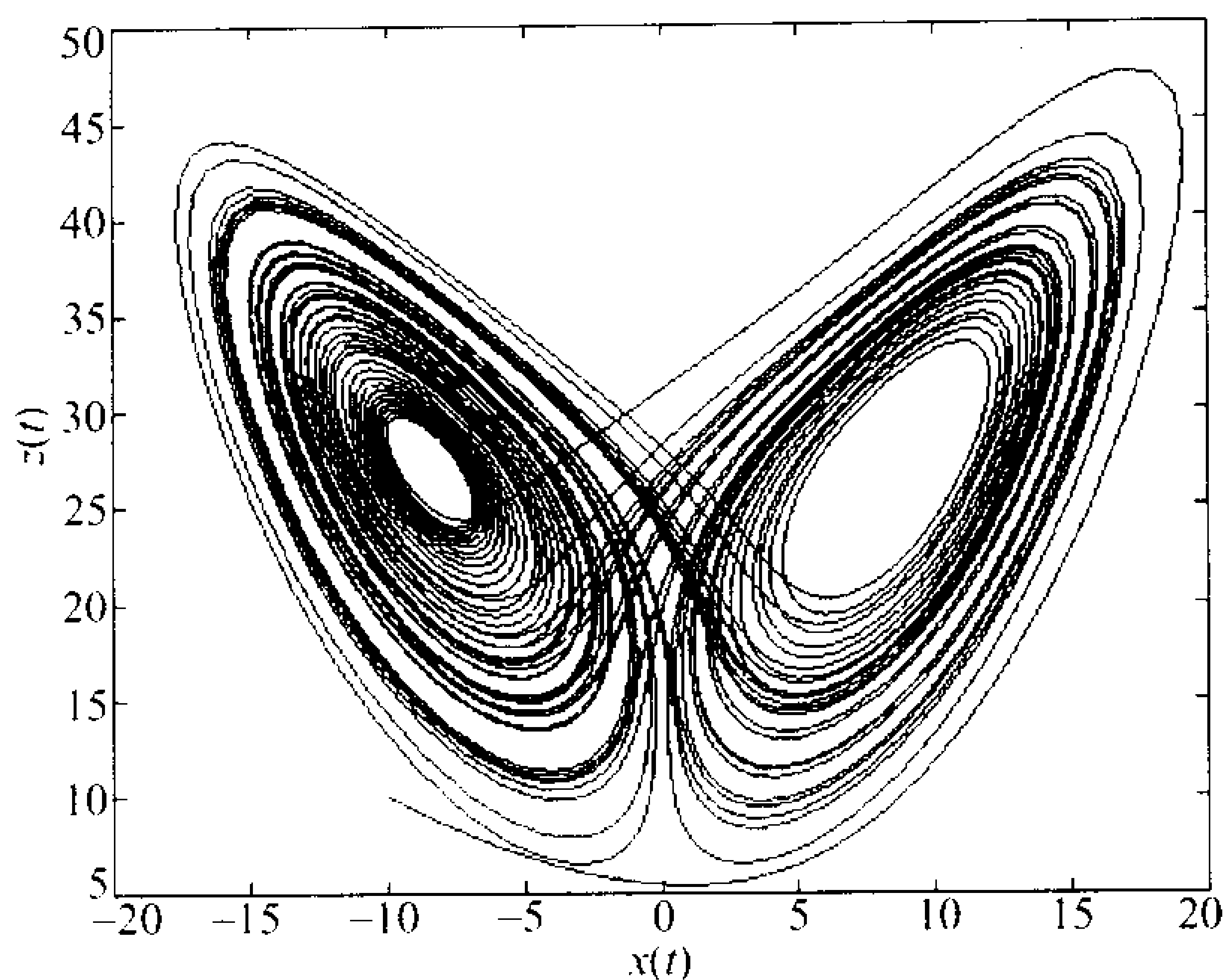


图 2.1.13 Lorenz 吸引子

Lorenz 吸引子的每一片都不是二维曲面，而是多层曲面，是三维空间中的分形曲面，其分数维 $D_C = 2.06$.

设 z 的第 n 次迭代的最大值为 z_n, 则相邻两次迭代值 z_{n+1} 与 z_n 的关系由图 2.1.14 所示.

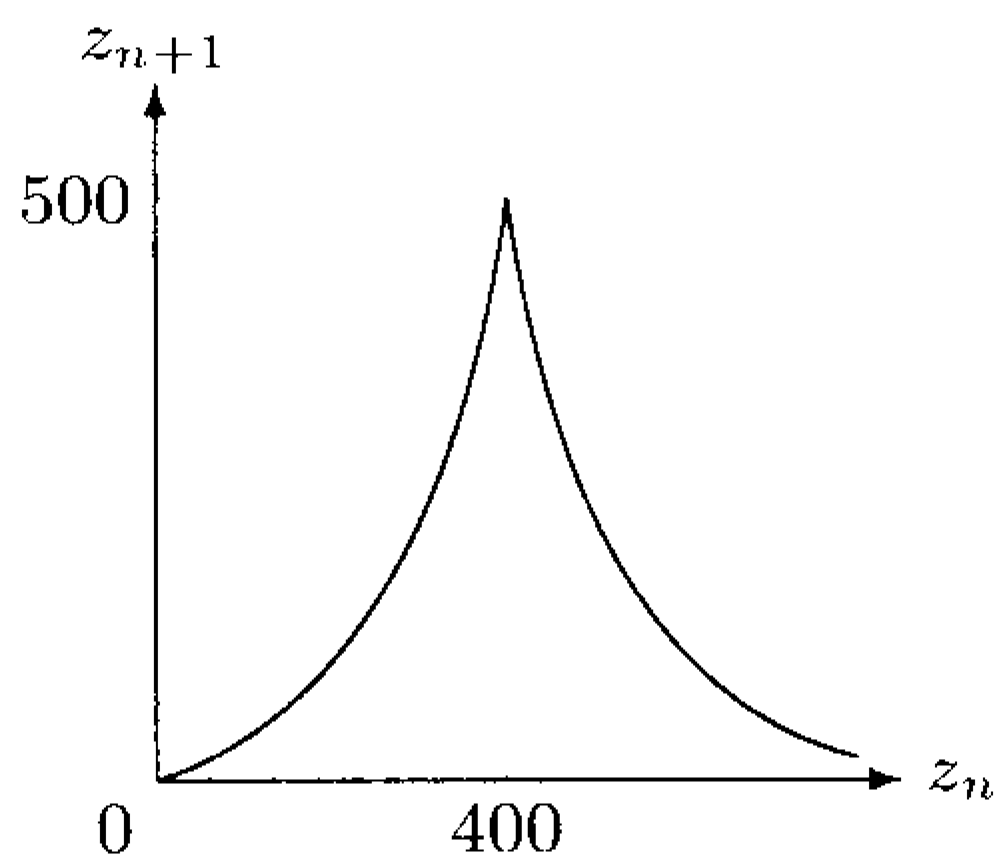


图 2.1.14 前 6000 次迭代的 z 值

由图 2.1.14 可见， z_{n+1} 与 z_n 的关系的图像非常类似于帐篷映射

$$T(y) = \begin{cases} 2y, & y \leqslant \frac{1}{2}, \\ 2 - 2y, & y > \frac{1}{2}, \end{cases}$$

$z_{n+1} = T(z_n)$.

若将 Lorenz 方程数值计算迭代 1400 次至 1900 次, 则轨线在 yz 平面上的投影如图 2.1.15 所示.

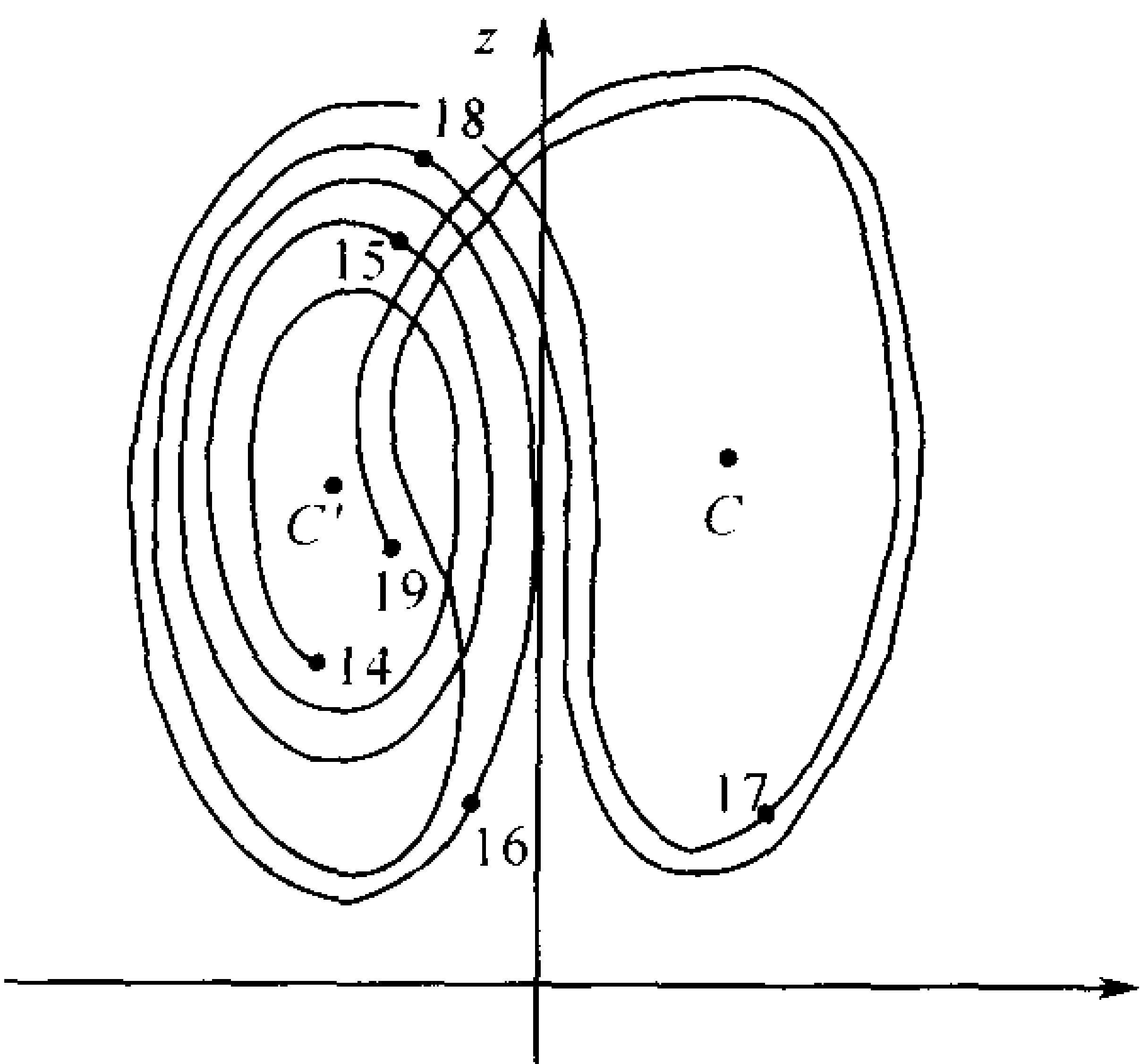


图 2.1.15 轨线在 yz 平面上的投影

2.1.4 时滞动力学模型与问题

在前面介绍的连续动力学模型中, 当前时刻 t 的状态变量的变化率仅与当前时刻 t 的状态变量有关, 但是实际上有许多事物的变化规律不仅依赖于当前时刻 t 的状态, 还依赖于历史上过去的状态. 换言之, 历史上的状态有 "记忆", 对当前的状态变化还起着作用. 而在动力学模型中, 代替常微分方程的是微分差分方程, 或者称为更复杂一些的具有偏差变元的微分方程、泛函微分方程. 从以下几个实例可见, 在实际问题中这类模型具有一定的广泛性.

例 2.1.9 一维的具时滞的 Logistic 方程

考虑诸如幼虫到成虫的成长期及一些动物的怀胎期的因素 (例如澳洲的袋鼠), 就要研究如下形式的方程:

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = rN(t)\left[1 - \frac{N(t-T)}{K}\right] \tag{2.1.32}$$

或

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = rN(t)\left[1 - \frac{1}{K}\int_{-\infty}^{t} W(t-s)N(s)\mathrm{d}s\right]. \tag{2.1.33}$$

系统 (2.1.32) 中的 $T>0$ 为离散的常时滞，即种群数量 $N(t)$ 的变化率不仅和现在时刻 t 的 $N(t)$ 有关，而且与在 $t-T$ 时刻的状态 $N(t-T)$ 有关，称之为**具有时滞的微分差分方程**. 系统 (2.1.33) 表明种群数量的变化率不仅和现在时刻 t 的 $N(t)$ 有关，而且与 t 时刻以前的整个一段时间的种群数量值 $N(s)$ $(-\infty < s \leqslant t)$ 有关，称之为 **具有时滞的微分积分方程**. 给这两种方程一个统一的名称为 **时滞的泛函微分方程**.

例 2.1.10 1973 年， London 和 Yorke 研究了麻疹传播的模型，其形式为

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = \beta(t)S(t)[S(t-12) - S(t-14) - 2r] + r, \tag{2.1.34}$$

其中 $S(t)$ 表示在时刻 t 无免疫力的个体数目， r 是这种个体在人口中所占的比例， $\beta(t)$ 为人口特征函数，常数滞量 14 与 12 的生物意义是潜伏期的上限和下限. 系统 (2.1.34) 是一种多个常时滞的微分差分方程.

例 2.1.11 在人口增长理论中，出现了所谓的 **除旧更新方程**

$$u(t) = f(t) + \int_{0}^{t} u(t-\tau)\frac{\mathrm{d}G(\tau)}{\mathrm{d}\tau}\mathrm{d}\tau. \tag{2.1.35}$$

例 2.1.12 在血液动力学及种群动力学中出现的非线性微分差分方程为

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \frac{AB^n x(t-T)}{B^n + x^n(t-T)} - Cx(t). \tag{2.1.36}$$

例 2.1.13 在火箭燃烧的控制理论中，钱学森、宋健在专著《工程控制论》中提出了如下形式的方程：

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} + (1-n)u(t) + nu(t-\tau) = 0. \tag{2.1.37}$$

例 2.1.14 在数理统计中，关于资本主义经济的周期性危机提出过如下形式的方程:

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = au(t) + bu(t-\tau) + f(t). \tag{2.1.38}$$

从上面的例子可见，无论在物理、力学、工程技术中或者在生物医学中，甚至在经济学研究中有大量的现象，若要精确地描述其整个动力学过程，则必须考虑“记忆”效应，即考虑时滞的因素，这是非常必要的. 当然也有不少现象把这种“记忆”效应略去了，即认为时滞非常小而接近于无时滞的情形. 此外对这种时滞动力学的研究手段及结果相当缺乏，也是造成不得不考虑无时滞情形的原因. 近几十年来，在数学上，时滞动力系统的研究有了较大的长进，已有了一些新的工具和方法，因此有必要使广大科技工作者了解这方面的进展以及应用的实例，本书将对这方面的内容作一些必要的介绍.

以下首先介绍这类系统的初值问题的提法以及有关的基本性质的结果，然后介绍这类系统的吸引子问题.

1. 具有界时滞的泛函微分方程的初值问题的基本性质

关于初值问题的提法，以一阶的常时滞的微分差分方程为例，其形式为

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(t, x(t), x(t-\tau)), \quad \tau > 0. \tag{2.1.39}$$

设初始条件给在区间 $[t_0-\tau, t_0]$ 上，即

$$x(t) = \varphi(t), \quad t \in [t_0-\tau, t_0], \tag{2.1.40}$$

其中 $\varphi(t)$ 为已知函数，它是关于自己的变元的连续函数，先求 $t_0 \leqslant t \leqslant t_0+\tau$ 时的 (2.1.39) 的解，实际上就是在 $[t_0, t_0+\tau]$ 上求下列常微分方程的初值问题的解:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(t, x(t), \varphi(t-\tau)), \tag{2.1.41}$$

$$x(t_0) = \varphi(t_0). \tag{2.1.42}$$

记上述初值问题 (2.1.41), (2.1.42) 在区间 $[t_0, t_0+\tau]$ 上的解为 $x=\varphi_1(t)$, 对于 $t\in[t_0, t_0+2\tau]$, $x(t-\tau)=\varphi_1(t-\tau)$, 于是方程 (2.1.39) 带有初始条件 $x(t_0+\tau)=\varphi_1(t_0+\tau)$ 的常微分方程的初值问题:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=f(t,x(t),\varphi_1(t-\tau)), \tag{2.1.43}$$

$$x(t_0+\tau)=\varphi_1(t_0+\tau) \tag{2.1.44}$$

在区间 $[t_0+\tau, t_0+2\tau]$ 的解存在, 记之为 $x(t)=\varphi_2(t)$. 依次类推, 在区间 $[t_0+n\tau, t_0+(n+1)\tau]$ 上求解如下的常微分方程的初值问题:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=f(t,x(t),\varphi_n(t-\tau)), \tag{2.1.45}$$

$$x(t_0+n\tau)=\varphi_n(t_0+n\tau), \tag{2.1.46}$$

其中 $x=\varphi_n(t)$ 是区间 $[t_0+(n-1)\tau, t_0+n\tau]$ 上的解. 以上的求解初值问题 (2.1.39), (2.1.40) 的方法称为 **分步法**.

从上面求解 (2.1.41), (2.1.42) 的过程易知, 若固定 $\varphi(t_0)$ 又改变在 $[t_0-\tau, t_0)$ 上的 $\varphi(t)$ 之值, 初值问题 (2.1.41), (2.1.42) 的解有许多不同的形式, 因此, 如果仍像常微分方程初值问题一样仅给出初始时刻 t_0 的条件 $x_0=\varphi(t_0)$, 那么这种解将是不确定的, 从而是无实际意义的.

对于一般的具时滞的微分差分方程以及微分积分方程可以用下面形式的具时滞的泛函微分方程统一起来:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=f(x,x_t). \tag{2.1.47}$$

设 $C([a,b],\mathrm{R}^n)$ 表示将区间 $[a,b]$ 映射到 R^n 中的连续向量函数并具有一致收敛拓扑的 Banach 空间. 一般地, 对给定的 $\tau>0$, 我们将空间 $C([-\tau,0],\mathrm{R}^n)$ 简记为 C. 对每一 $\varphi\in C$, 其范数定义为 $\|\varphi\|=\sup\limits_{-\tau\leqslant\theta\leqslant 0}|\varphi(\theta)|$, 其中 $|\cdot|$ 是 R^n 中的范数. 定义 $x_t\in C$ 为 $x_t(\theta)=x(t+\theta)$, θ 取遍 $[-\tau,0]$ 上的一切值. 设 $D\subset\mathrm{R}\times C$, 即 $(t,x_t)\in D$, 而 f 为泛函, 定义为 $f: D\to\mathrm{R}^n$, 实际上 (2.1.47) 严

格讲是具有界滞量的时滞型泛函微分方程，其中 $\frac{\mathrm{d}x(t)}{\mathrm{d}t}$ 表示 $x(t)$ 对 t 的右导数.

方程 (2.1.47) 的初值条件为

$$x_{t_0}=\varphi,\quad \varphi\in C \text{ 给定}. \tag{2.1.48}$$

实际上 $x_{t_0}(\theta)=x(t_0+\theta)=\varphi(\theta)$, $-\tau\leqslant\theta\leqslant 0$.

初值问题 (2.1.47), (2.1.48) 的解 $x(t;t_0,\varphi)$ 是指存在 $A>0$, 使得 $x(t;t_0,\varphi)$ 是方程 (2.1.47) 在区间 $[t_0-\tau,t_0+A)$ 上的解, 且 $x_{t_0}(\theta;t_0,\varphi)=x(t_0+\theta;t_0,\varphi)=\varphi(\theta)$, $-\tau\leqslant\theta\leqslant 0$. 有时也称 $x(t;t_0,\varphi)$ 是方程 (2.1.47) 过点 (t_0,φ) 的解.

易见 (2.1.47) 是一种相当广泛的方程，它包含了常微分方程 $\dot{x}(t)=f(t,x(t))$, 因为当 $\tau=0$ 时， C 空间成为 R^n 空间， x_t 成为 $x(t)$, $f(t,x_t)$ 实际上成为 $f(t,x(t))$ 了.

又 (2.1.47) 包含了滞量为有界的时滞型微分差分方程组. 例如

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=F(t,x(t),x(t-\tau_1),\cdots,x(t-\tau_n)), \tag{2.1.49}$$

其中 F 为 n 维向量值函数，又 $\tau_i\leqslant\tau$ $(i=1,2,\cdots,m)$, (2.1.49) 要记成 (2.1.47) 的形式，只要取

$$f(t,\psi)=F(t,\psi(0),\psi(-\tau_1),\psi(-\tau_2),\cdots,\psi(-\tau_n)),\quad x\in C,$$

把 ψ 取为 x_t, 则 (2.1.49) 就变成 (2.1.47) 的形式.

方程 (2.1.47) 还包括如下形式的微分积分方程:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=\int_{-\tau}^{0}g(t,\theta,x(t+\theta))\mathrm{d}\theta. \tag{2.1.50}$$

当 (2.1.47) 的右边 $f(t,\psi)=\displaystyle\int_{-\tau}^{0}g(t,\theta,\psi(\theta))\mathrm{d}\theta$ 时，取 $\psi=x_t$, 则 (2.1.50) 就写成 (2.1.47) 形式了.

从形式上看，泛函微分方程 $\dot{x}=f(t,x_t)$ 与常微分方程 $\dot{x}(t)=f(t,x(t))$ 很类似，其区别只是前者的 f 定义在 $\mathrm{R}\times C$ 空间，而后

者的 f 定义在 $\mathrm{R}\times\mathrm{R}^n$ 空间. 因此, 常微分方程中的许多理论都可以平行推广到泛函微分方程中来. 另一方面, 我们必须注意到 C 空间是无限维的, 它不具备有限维空间 R^n 那么多良好的性质. 例如 R^n 空间中的有界闭集与紧集是等价的, 但 C 空间中却不是这样. 因此, 常微分方程中的许多性质在泛函微分方程中是没有的. 当然, 也有一些泛函微分方程中有的性质, 在常微分方程中未必具有, 但是在以往不少的科技论文中常把具时滞的微分差分方程的右端按 τ Taylor 展开, 例如把方程 (2.1.39) 的右端按 τ 展开成

$$f(t,x(t),x(t-\tau)) = f(t,x(t),x(t)) - \left.\frac{\partial f(t,x(t),v)}{\partial v}\right|_{v=x(t)} x'(t)\cdot\tau + o(\tau^2),$$

然后用常微分方程

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(t,x(t),x(t))$$

作为 (2.1.39) 的近似. 这里的操作的可行性条件是要求 τ 相当小, 但事实上许多实际问题中 τ 是有一定大小的, 不能用 $\tau=0$ 代替存在 τ 的情形. 对有时滞的动力学系统的研究一定要强调滞量 τ 的作用, 分清楚有滞量 τ 与无滞量 τ (即 $\tau=0$) 的区别或者大的滞量 τ 与小的滞量 τ 的区别.

有关泛函微分方程的初值问题 (2.1.47), (2.1.48) 的解的基本性质有:

(1) 如果方程 (2.1.47) 的右端 f 是连续的, 那么存在 $A>0$, 使得 (2.1.47), (2.1.48) 的解在 $[t_0-\tau, t_0+A)$ 上存在. 要注意的是解仅在 $[t_0, t_0+A)$ 即 t_0 的右边存在, 并不像常微分方程的初值问题解在初始时刻 t_0 的两边的适当区间均存在.

(2) 如果 f 是连续且满足 Lipschitz 条件

$$\|f(t,\phi) - f(t,\psi)\| \leqslant L\|\phi-\psi\|,$$

那么 (2.1.47), (2.1.48) 在区间 $[t_0-\tau, t_0+A)$ 上的解是唯一的. 要注意的是此时虽然两个解的轨线可以相交, 甚至相交无穷多次, 但两

个解 ψ_1, ψ_2 作为 C 空间中的两个元素是不同的, 即 $\|\psi_1-\psi_2\|_C \neq 0$.

例 2.1.15 考虑下面两个初值问题

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -x\left(t-\frac{\pi}{2}\right), \tag{2.1.51}$$

$$x\left(\frac{\pi}{4}+\theta\right) = \sin\left(\theta+\frac{\pi}{4}\right), \quad \theta \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right] \tag{2.1.52}$$

以及

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -x\left(t-\frac{\pi}{2}\right), \tag{2.1.53}$$

$$x\left(\frac{\pi}{4}+\theta\right) = \cos\left(\theta+\frac{\pi}{4}\right), \quad \theta \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right]. \tag{2.1.54}$$

实际上 $x(t)=\sin t$ 与 $x(t)=\cos t$ 分别为上述两个初值问题的解. 显见 (2.1.51) 和 (2.1.53) 的右端满足解的存在唯一性定理的条件. 但是 $\sin t$ 与 $\cos t$ 在 O-tx 平面中的轨线图上相交无限次, 但这并不破坏上面两个初值问题的解的唯一性, 因为我们所说的解的唯一性是指在 $C([-\frac{\pi}{2},0],\mathrm{R})$ 中的两个元素即两个解 ψ_1 与 ψ_2 是不相同的, 即 $\|\psi_1-\psi_2\|_{C([-\frac{\pi}{2},0],\mathrm{R})} \neq 0$, 也就是说 $\sin t$ 与 $\cos t$ 在 $C([-\frac{\pi}{2},0],\mathrm{R})$ 中是不同的.

(3) 解只能向前延展 $(t>t_0)$, 而向后 $(t, t_0-\tau)$ 延展的解可能不存在, 这也是时滞型泛函微分方程与常微分方程的不同之处, 见下例.

例 2.1.16 考虑纯量方程

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = ax(t)+bx(t-\tau), \tag{2.1.55}$$

其中 a, b, τ 皆为常数, $b\neq 0, a\neq -b, \tau>0$.

现取初始时刻 $t_0=0$, 初始函数 $\varphi(\theta)\equiv k$ (常数), $\theta\in[-\tau,0]$. 将 (2.1.55) 写成

$$x(t-\tau) = \frac{\dfrac{\mathrm{d}x(t)}{\mathrm{d}t}-ax(t)}{b}.$$

令 $s=t-\tau$, 则有

$$x(s)=\frac{\dot{x}(s+\tau)-ax(s+\tau)}{b}.$$

当 $s\in[-2\tau,-\tau]$ 时, $x(s+\tau)=\varphi(s+\tau)=k$, $\dot{x}(s+\tau)=0$, 故 $x(s)=-\dfrac{ak}{b}$. 另一方面, 由初始条件知 $x(-\tau)=\varphi(-\tau)=k$, 由于 $a\neq -b$, 即 $-\dfrac{ak}{b}\neq k$, 所以 $x(t)$ 在 $-\tau$ 处不连续. 故方程 (2.1.55) 不存在满足初值 $\varphi(\theta)\equiv k$, $\theta\in[-\tau,0]$ 的反向延展的解.

2. 具有界时滞的泛函微分方程的初值问题解的渐近性态, 吸引子问题

考虑系统 (2.1.47) 的解的渐近性态, 与连续的常微分方程系统一样, 具有 3 种解的渐近性态: 平衡点、极限环、奇异吸引子.

平衡点即常值函数向量 $x_t=a$, 它满足 $f(t,a)=0$. 一般设平衡点位于原点, 即 $x_t=0$, 它满足 $f(t,0)=0$. 关于平衡点 $x=0$ 的稳定性, 泛函微分方程与常微分方程有一个比较大的差别, 看下面的例子.

例 2.1.17 考虑如下纯量方程

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=b(t)x\left(t-\frac{3\pi}{2}\right), \tag{2.1.56}$$

其中

$$b(t)=\begin{cases} g(t), & t\leqslant 0, \\ 0, & 0\leqslant t\leqslant \dfrac{3\pi}{2}, \\ -\cos t, & \dfrac{3\pi}{2}\leqslant t\leqslant 3\pi, \\ 1, & t\geqslant 3\pi, \end{cases} \tag{2.1.57}$$

$g(t)$ 为连续函数, 当初始时刻 t_0 及初值函数 φ 取如下形式 $t_0=0$,

$\varphi \in ([-\frac{3\pi}{2},0],\mathrm{R})$ 时有解

$$x(t)=\begin{cases}\varphi(0), & 0\leqslant t\leqslant \dfrac{3\pi}{2},\\[2ex] (-\sin t)\varphi(0), & t\geqslant \dfrac{3\pi}{2}.\end{cases}$$

易见当 $\|\varphi\|_C$ 很小时，即 $|\varphi(0)|$ 很小时， $\|x(t)\|$ 也很小，故零解对初始时刻 $t_0=0$ 是稳定的. 但是，如果取 $t_1>3\pi$, 将 t_1 作为初始时刻，则对于 $t\geqslant t_1$ 时，方程为

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=x\left(t-\frac{3\pi}{2}\right), \tag{2.1.58}$$

它有解为

$$x(t)=a\mathrm{e}^{\lambda_0 t}, \tag{2.1.59}$$

其中 a 为任意常数， $\lambda_0>0$ 满足方程 $\lambda=\mathrm{e}^{-\frac{3\pi}{2}\lambda}$. 若取 $\psi(\theta)\equiv\widetilde{a}$, $\theta\in[-\frac{3\pi}{2},0]$, $\psi\in C([-\frac{3\pi}{2},0],\mathrm{R})$, 初始条件为 $x_{t_1}=\psi$, $x(t_1+\theta)=\psi(\theta)$, $\theta\in[-\frac{3\pi}{2}+t_1,t_1]$, 则 (2.1.58) 的解为

$$x(t)=\widetilde{a}\mathrm{e}^{\lambda_0(t-t_1)}\quad (t\geqslant t_1).$$

虽然 $|\widetilde{a}|$ 可以取为任意小，但是总存在某时刻 T, 使 $|x(T)|$ 适当大，即对初始时刻是不稳定的. 这种现象对实际应用是很不方便的, 所以我们只好加强稳定性的定义, 即要求零解的稳定等价于每一时刻的零解都稳定. 显见在通过解的表达式验证时也要比常微分方程的情形复杂得多, 但是有趣的是在将 Lyapunov 函数方法推广为 Lyapunov 泛函方法后，本质上并没有带来多大的复杂性.

关于极限环以及奇异吸引子，由于有时滞的泛函微分方程本质上是属于无限维的动力学系统，因此出现极限环以及奇异吸引子是比较容易的，譬如常微分方程自治系统只有三维以上的系统才有可能出现混沌奇异吸引子. 但是即使有时滞的泛函微分数量方程, 即一维的模型也有可能出现奇异吸引子. 这些领域的研究刚刚开始，本章在 §2.2 中将有简单介绍.

2.1.5 时空动力学模型与问题

用时空动力学模型描述的实际问题, 包括物理力学、生物学、信息科学、化学等自然科学的许多问题. 在这些系统中, 状态不仅随时间的变化而变化, 而且还与空间所有点的位置有关. 得到的时空动力学模型可以用偏微分方程描述, 以反应扩散过程为例, 可以得到反应扩散方程

$$\frac{\partial u}{\partial t} = D^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) + F(u). \tag{2.1.60}$$

例 2.1.18 耗散系统的三分子模型由 Prigogine 和 Lefever 在 1968 年提出, 1973 年由 Tyson 命名为 Brussel 振子. 它描述系统中某些化学反应的中间产物的浓度随状态变量 x 和 y 演化过程的模型, 如果考虑浓度不均匀有扩散的情形, 可以归结为如下模型

$$\frac{\partial x}{\partial t} = A - (B+1)x + x^2 y + D_x \Delta x, \tag{2.1.61}$$

$$\frac{\partial y}{\partial t} = Bx - x^2 y + D_y \Delta y, \tag{2.1.62}$$

这里 $x(x_1, x_2, x_3, t)$, $y(x_1, x_2, x_3, t)$ 为两个状态变量, D_x 和 D_y 分别为 x 和 y 的扩散系数, Δ 记号为 Laplace 算子

$$\Delta x = \frac{\partial^2 x}{\partial x_1^2} + \frac{\partial^2 x}{\partial x_2^2} + \frac{\partial^2 x}{\partial x_3^2},$$

$$\Delta y = \frac{\partial^2 y}{\partial x_1^2} + \frac{\partial^2 y}{\partial x_2^2} + \frac{\partial^2 y}{\partial x_3^2}.$$

例 2.1.19 种群生态学中的种群密度分布不均匀的模型. 种群密度假定为 $N_1(x, y, z, t)$, $N_2(x, y, z, t)$, 若高密度位置的种群要向低密度位置扩散, 则有

$$\frac{\partial N_1}{\partial t} = \Delta N_1 + N_1 F_1(N_1, N_2), \tag{2.1.63}$$

$$\frac{\partial N_2}{\partial t} = \Delta N_2 + N_2 F_2(N_1, N_2), \tag{2.1.64}$$

其中 $\Delta N_i = \frac{\partial^2 N_i}{\partial x^2} + \frac{\partial^2 N_i}{\partial y^2} + \frac{\partial^2 N_i}{\partial z^2}$.

例 2.1.20 神经传导的 Fitz–Hugh–Nagumo 方程组

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + u(1-u)(u-a) - w, \tag{2.1.65}$$

$$\frac{\partial w}{\partial t} = bu - \gamma w, \tag{2.1.66}$$

其中 a, b, γ 均为正常数且 $0 < a < 1$.

假如对于时空动力学系统中的空间变量、时间变量以及状态变量 3 种变量的取值分别考虑为离散值与连续值的状况，那么其动力学模型可以列于表 2.1.2.

表 2.1.2 各类时空动力学模型及特征

动力学模型	空间变量	时间变量	状态变量	演化方式
偏微分方程	连续	连续	连续	连续
元胞自动机	离散	离散	离散	离散同步
耦合映像格子机	离散	离散	连续	离散同步
M–P 自动机 (离散神经网络)	离散	离散	离散或连续	离散同步或异步 (有序)

下面以实例说明除了偏微分方程以外的另外 3 种模型.

例 2.1.21 考虑如下简单的 $r = 1$, $k = 2$ 的一维元胞自动机，其中 r 为随时间演化方式的空间变量的半径， k 为状态变量仅取离散值的个数，它的演化规则为

$$S_i(t+1) = f(S_{i-1}(t), S_i(t), S_{i+1}(t)) \quad (i = 1, 2, \cdots, N). \tag{2.1.67}$$

一般地 f 的种类数目为 $k^{k^{2r+1}}$，本例为 $2^{2^3} = 2^8 = 256$.

对于上述不同的演化生成规则，整个演化过程的终态有三种情形：不动点、周期点、混沌奇异吸引子及具自组织性的破碎性的其它复杂吸引子.

实际上，3 个相邻的格点，每个格点上取值 0 或 1, 相当于二进制数取 3 位，共有 8 种数，分别记为

十进制数	7	6	5	4	3	2	1	0
二进制数	111	110	101	100	011	010	001	000

如果某种规则 f 使上述数对应 0 或 1, 又按二进制数排列，记它相应的十进制数为演化规则的编号，例如二进制数 01011010 为十进制数 90, 称之为第 90 号生成规则. 规则 90 可以用表 2.1.3 来表示输入与输出的对应关系：

表 2.1.3　规则 90 的输入输出对应关系

输入状态	111	110	101	100	011	010	001	000
输入	0	1	0	1	1	0	1	0

可以发现随着时间的向前推进，上述 256 种演化规则总体上状态的演化有下述三种情况：

(1) 从任何初始条件出发，状态演化最终收敛于某对应时间空间的不动点. 例如规则 128 (10000000) 最终状态全为 0 以及规则 250 (11111010) 最终状态全为 1 (见图 2.1.16).

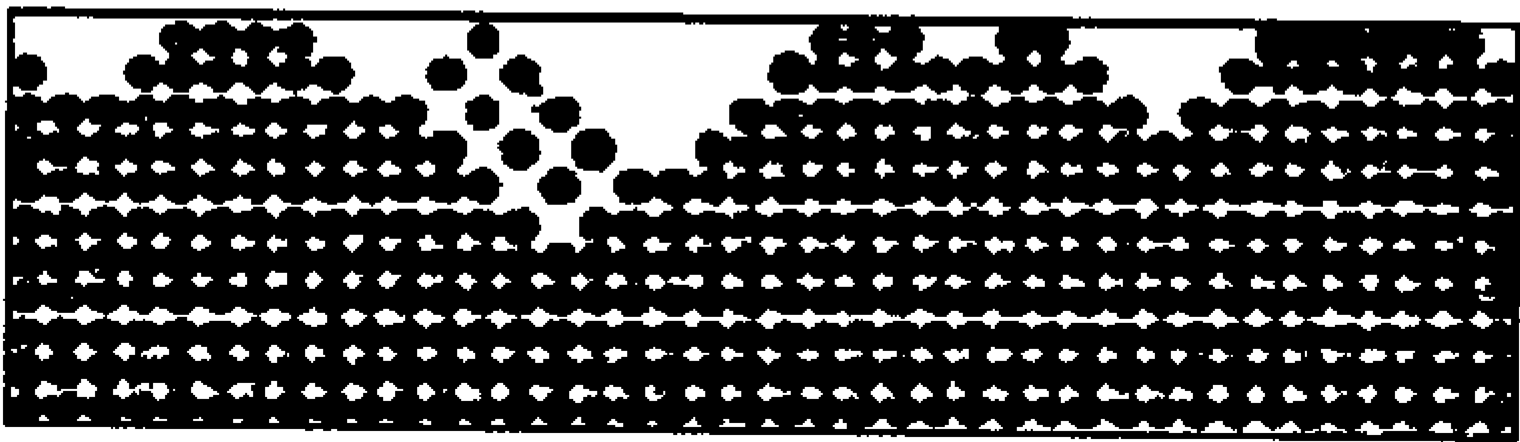


图 2.1.16　规则 250 (11111010)

(2) 对应时间演化的最终极限状态为极限环，即一定的空间区

域随时间的变化是周期的，例如规则 178 (10110010) (见图 2.1.17)，规则 108 (01101100) (见图 2.1.18).

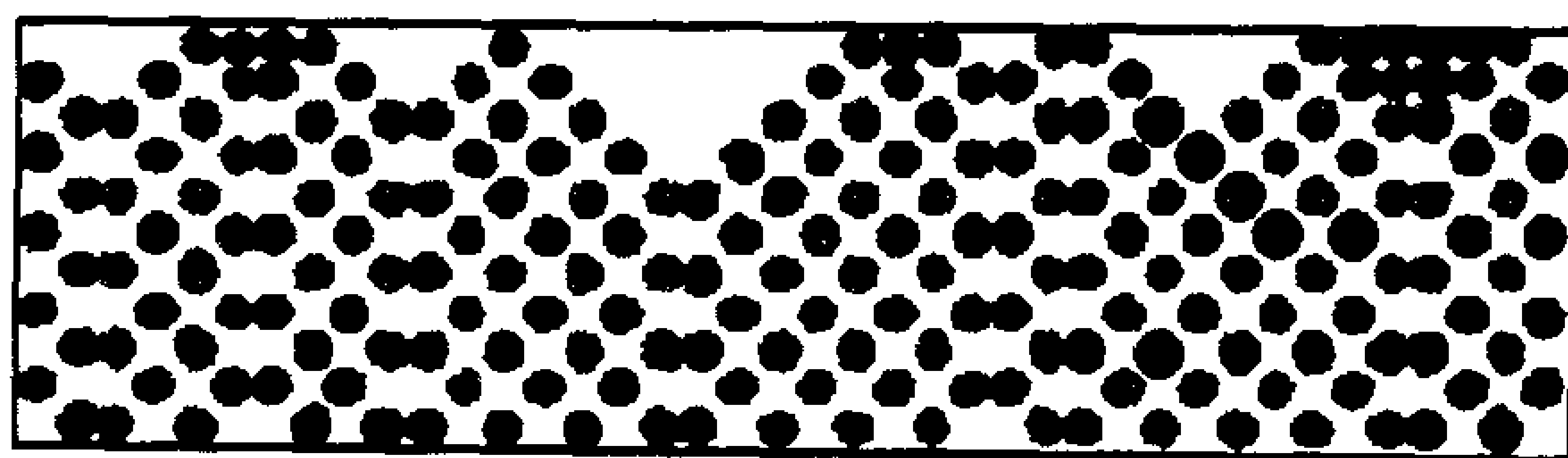


图 2.1.17　规则 178 (10110010)

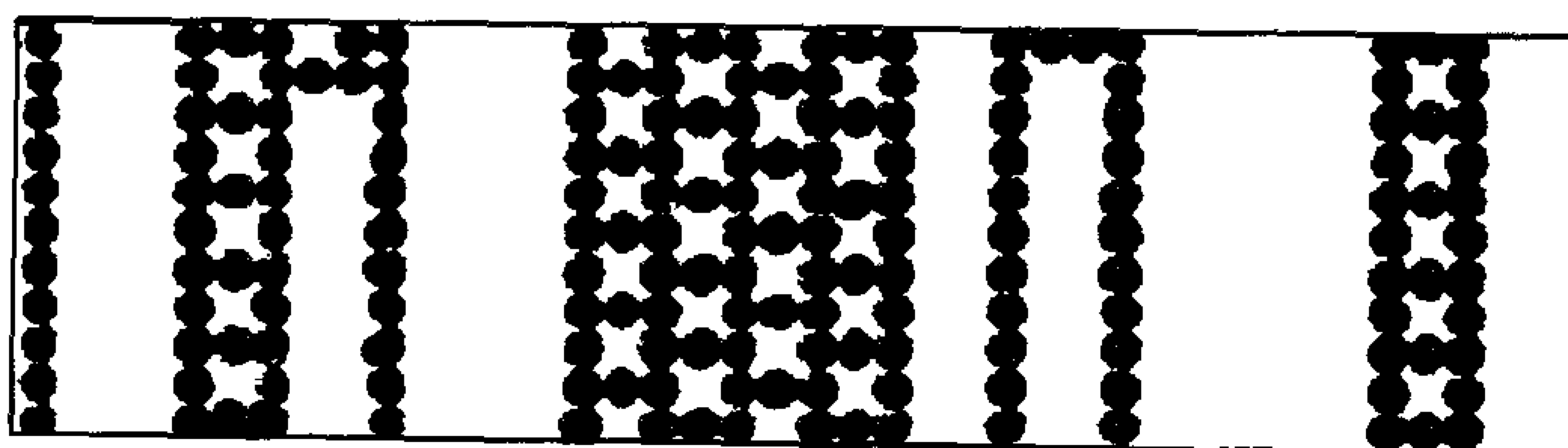


图 2.1.18　规则 108 (01101100)

(3) 时间空间演化的最终状态为最终极限状态是非周期性的图式呈现出分形的或破碎的、具自组织性的混沌奇异吸引子与其它类型的复杂的吸引子，例如规则 90 (01011010) (见图 2.1.19)，规则 126 (01111110) (见图 2.1.20).

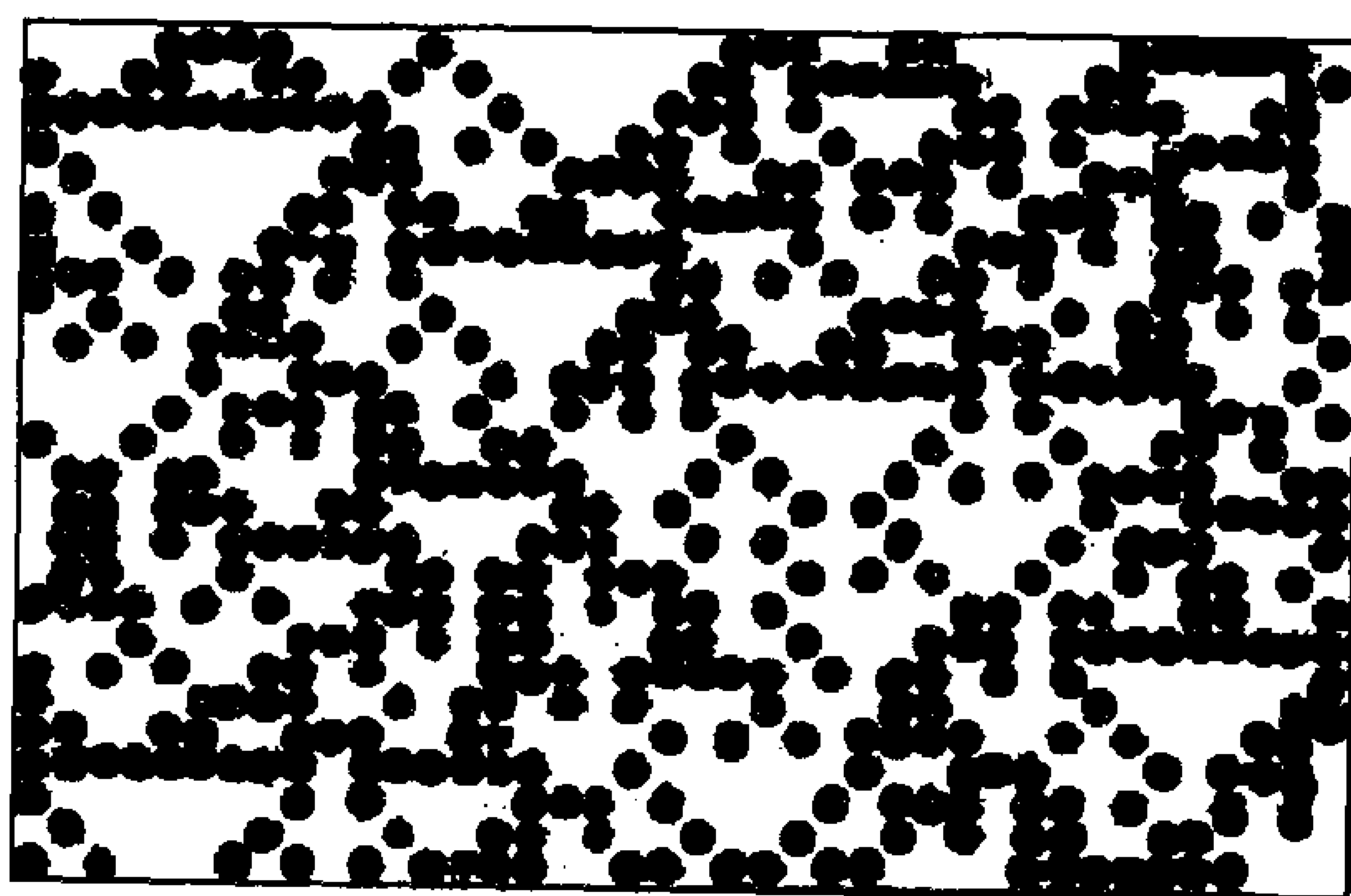


图 2.1.19　规则 90 (01011010)

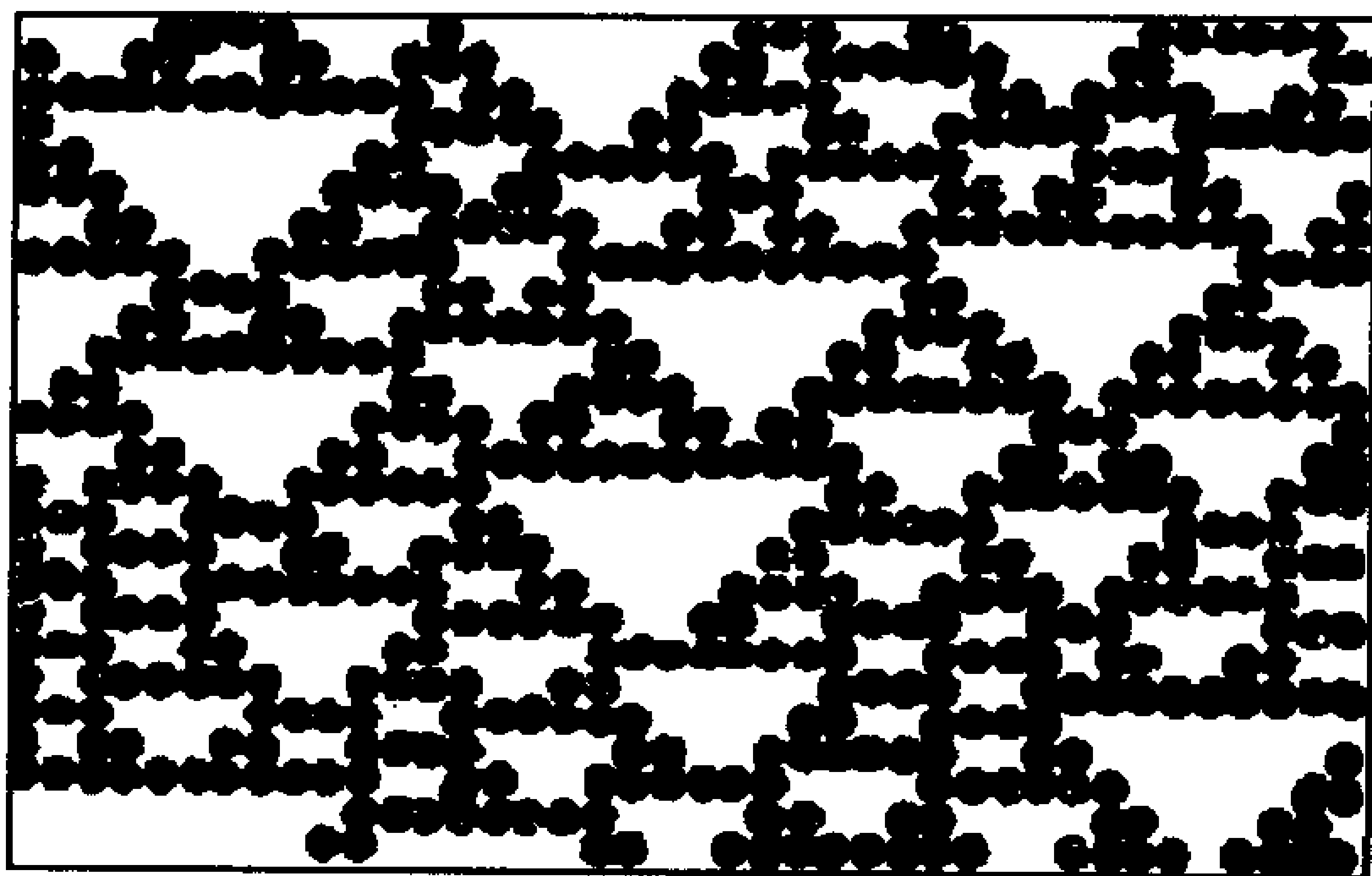


图 2.1.20　规则 126 (01111110)

例 2.1.22　将一维反应扩散方程

$$\frac{\partial u(x,t)}{\partial t}=\varepsilon\frac{\partial^2 u(x,t)}{\partial x^2}+f(u(x,t)),\quad 0\leqslant x\leqslant L \tag{2.1.68}$$

及周期性的边界条件

$$u(0,t)=u(L,t) \tag{2.1.69}$$

按照时间变量离散、空间变量离散、状态变量保持连续的方法构造一种新的系统，称之为 **耦合映像格子系统**.

将 (2.1.68) 的局部反应过程用一个非线性映像来表述, 将状态变量改为 $x_n(i)$, 其中 n 表示时间离散化坐标， i 代表空间离散化坐标, 即 i 为格点坐标 $(i=1,2,\cdots,L)$. 关于扩散过程则将 Laplace 算子离散化，得到如下模型

$$x'_n(i)=f(x_n(i)), \tag{2.1.70}$$

$$x_{n+1}(i)=(1-\varepsilon)f(x_n(i))+\frac{\varepsilon}{2}[f(x_n(i-1))+f(x_n(i+1))], \tag{2.1.71}$$

$$x_n(0)=x_n(L),\quad i=1,2,\cdots,L. \tag{2.1.72}$$

特别地取 $f(x)=1-ax^2$, 即单峰映像，此时人们称之为 **耦合单峰格子** (coupled logistic lattices). 注意这个模型中由第 n 步时间的状态变量生成第 $n+1$ 步时间的状态变量是以同步的方式进行的.

这种模型的想法，最初由 K. Kanoko 等提出，近年来引起了人们的广泛注意. 他得到了耦合单峰格子的参数 (a,ε) 空间相图，在图 2.1.21 中用 I, II, III, IV, V, VI 分别表示六种相模式所占据的参数区域：

(I) 冻结化随机图案模式 (frozen random pattern);

(II) 图案选择模式 (pattern selection);

(III) 缺陷混沌扩散模式 (defect chaotic diffusion);

(IV) 缺陷湍流模式 (defect turbulence);

(V) 图案竞争阵发混沌模式 (pattern competition intermittency);

(VI) 完全发展湍流模式 (fully developed turbulence).

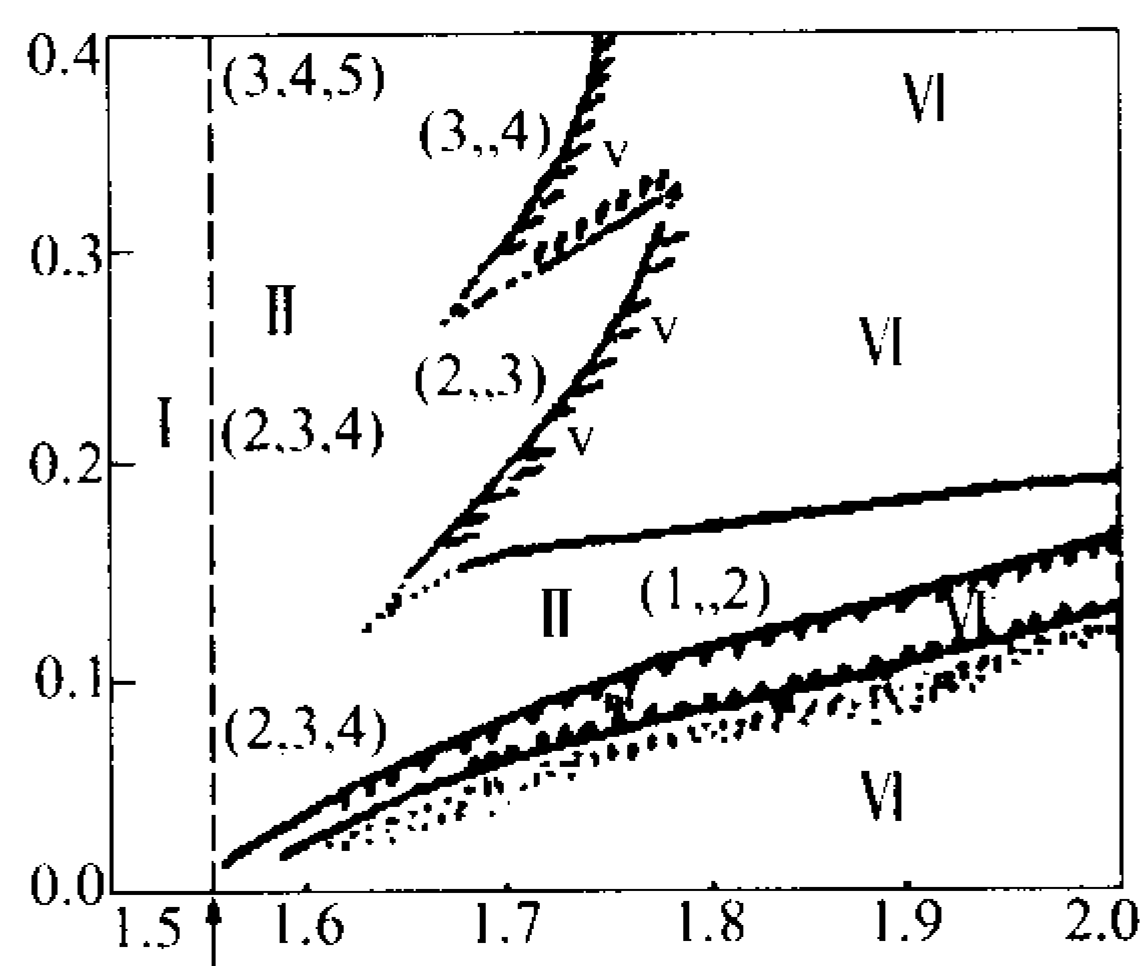


图 2.1.21　耦合单峰格子的参数空间相图

就一维耦合映像格子来说，耦合单峰格子的许多时空行为还是有较大的普遍性的，并且目前研究得相对比较深入，可以此作为研究一维耦合映像格子模型的范式.

例 2.1.23　如果在上面的例子中考虑系统中有长程相互作用或全局的耦合效应，又不是从反应扩散的偏微分方程的离散化来构造新的模型，而是直接从神经网络生物信息处理系统中一个神经元通过突触连接而传递信息的生物背景来构造全程耦合映像模

型，那么可以得到如下简化的神经网络模型：

$$y_{n+1}(i) = ky_n(i) + \sum_{j=1}^{N} w_{ij} f(y_n(j)) - \alpha f(y_n(i)) + a_i, \qquad (2.1.73)$$

这里状态变量 $y_n(i)$ 为在离散时间 n 时第 i 个神经元的内部状态，w_{ij} 为第 i 个神经元与第 j 个神经元的连接权值，N 为神经元总数，k, α 及 a_i 为常数，f 为单调的取值在 $[0,1]$ 的非线性函数响应，如取为

$$f(y) = \frac{1}{1+\mathrm{e}^{-\frac{y}{\varepsilon}}} \quad (\varepsilon > 0), \qquad (2.1.74)$$

而第 i 个神经元在离散时间为 n 时的输出，记为 $x_n(i)$, 它与 $y_n(i)$ 的关系为

$$x_n(i) = f(y_n(i)). \qquad (2.1.75)$$

模型 (2.1.73) 的推导以及它的复杂动力学行为将在后面介绍.

§2.2　非线性动力学分析与方法

非线性动力学研究的内容包括动力学演化过程、各种动力学模型的解的解析形式与数值近似值的求取、状态变量随时间演化的终态、极限状态即吸引子问题, 我们在这一节中着重于吸引子的分析与方法.

平衡点或不动点吸引子, 当讨论某个平衡点的局部稳定性时, Lyapunov 直接方法是一种有效的分析方法，它由 Lyapunov 基本定理作为理论依据, 但是对 Lyapunov 函数或泛函的符号有一定的限制. 这方面的内容将在 2.2.1 中介绍.

当考虑平衡点集合的整体稳定性时, 包括考虑周期解 (极限环) 的集合的稳定性分析时， Lyapunov 方法的推广 ——Lasalle 不变性原理是一个十分有效的分析方法，我们将在 2.2.2 中介绍. 对于混沌奇异吸引子的分析以及混沌动力学方法的介绍，将放在 2.2.3 中. 这方面涉及的内容很多, 我们只能着重于在混沌神经网络的分析中直接有应用的部分给予简单的介绍.

2.2.1 平衡点及 Lyapunov 稳定性分析

这里着重介绍常微分方程模型的 Lyapunov 函数方法与时滞型泛函微分方程模型的 Lyapunov 泛函的方法.

1. 常微分方程模型与 Lyapunov 函数方法

对于一般的 n 维连续动力学系统，即常微分方程模型

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x), \tag{2.2.1}$$

$x \in \mathrm{R}^n$, $f: \mathrm{R}^n \to \mathrm{R}^n$. 若 $f(0) = 0$, 即 0 是 (2.2.1) 的平衡点，以下给出零解 $x = 0$ 是稳定的、渐近稳定的与不稳定的定义.

定义 2.2.1 如果对于任意的 $\varepsilon > 0$ 及 $t_0 \in \mathrm{R}^+$, 存在 $\delta(t_0, \varepsilon) > 0$, 使得当 $\|x_0\| < \delta(t_0, \varepsilon)$ 时，方程组 (2.2.1) 以 (t_0, x_0) 为初值的解 $x(t; t_0, x_0)$ 在 $[t_0, +\infty)$ 上存在，且满足

$$\|x(t; t_0, x_0)\| < \varepsilon \quad (\forall t \in [t_0, +\infty)),$$

则称方程组 (2.2.1) 的零解 $x = 0$ 是 **稳定** 的. 即如果

$$\lim_{\|x_0\| \to 0} x(t; t_0, x_0) = 0$$

在 $[t_0, +\infty)$ 上关于 t 一致成立，那么零解 $x = 0$ 是稳定的.

定义 2.2.2 如果方程组 (2.2.1) 的解满足

(1) 零解 $x = 0$ 是稳定的;

(2) 存在 $\eta > 0$, 使得当 $\|x_0\| < \eta$ 时，有

$$\lim_{t \to +\infty} x(t; t_0, x_0) = 0,$$

则称 (2.2.1) 的零解 $x = 0$ 是 **渐近稳定** 的.

定义 2.2.3 如果存在 $\varepsilon_0 > 0$, 不管 $\delta > 0$ 怎样小，总存在 x_0^*, 虽然 $\|x_0^*\| < \delta$, 但是方程组 (2.2.1) 以 (t_0, x_0^*) 为初值的解 $x(t; t_0, x_0^*)$ 在 $[t_0, +\infty)$ 上不存在，或者存在 $t^* > t_0$, 使得

$$\|x(t^*; t_0, x_0^*)\| \geqslant \varepsilon_0,$$

则称方程组 (2.2.1) 的零解 $x=0$ 是 **不稳定** 的.

如果 (2.2.1) 系统有平衡点 $x=a$, 则可以作变换

$$y = x - a,$$

此时 (2.2.1) 化为

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(y+a). \tag{2.2.2}$$

因为 $f(a)=0$, 所以 (2.2.2) 具有零解, 而 (2.2.2) 的零解的稳定性对应 (2.2.1) 的平衡点的稳定性.

关于平衡点或零解的稳定性判别有两条途径.

第一条途径: 对于线性系统可以通过系数矩阵的特征值的实部来判别零解的稳定性; 而对于非线性系统 (2.2.1) 可以将之线性化, 然后根据一次近似定理将 (2.2.1) 的零解稳定性与线性化系统的零解稳定性对应起来. 注意这种对应关系对于不稳定及稳定情形有一些例外.

定理 2.2.1 (线性系统的稳定性) (1) 如果线性系统的系数矩阵的所有特征值的实部均为负的, 那么该线性系统的零解是渐近稳定的, 反之亦然.

(2) 如果线性系统的系数矩阵存在某个特征值具正实部, 那么该线性系统的零解是不稳定的.

(3) 假定线性系统的系数矩阵的所有特征值的实部均非正, 如果对应于实部为零的特征值的初等因子是单重的, 即该特征值是单重的, 那么该线性系统的零解是稳定的, 但不是渐近稳定的. 如果对应于实部为零的特征值的初等因子是多重的, 即该特征值是多重的, 那么该线性系统的零解是不稳定的.

定理 2.2.2 (一次近似定理) 设系统 (2.2.1) 中函数 f 是连续的, 且满足

(1) 如果一次近似组的零解是渐近稳定的, 那么原非线性系统 (2.2.1) 的零解也是渐近稳定的.

(2) 如果一次近似组的特征值中至少有一个为正实部，那么原非线性系统 (2.2.1) 的零解是不稳定的.

关于线性化系统的系数矩阵判别其特征值全为负实部的判据，常用的是 Routh–Hurwitz 判据. 假设特征方程有形式

$$g(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0, \tag{2.2.3}$$

构造如下形式的 $n \times n$ 的 Hurwitz 矩阵

$$H \triangleq \begin{pmatrix} a_1 & 1 & 0 & & & & & \\ a_3 & a_2 & a_1 & 1 & & & 0 & \\ a_5 & a_4 & a_3 & a_2 & a_1 & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & a_{n-1} & a_{n-2} \\ 0 & 0 & \cdots & \cdots & \cdots & \cdots & 0 & a_n \end{pmatrix}, \tag{2.2.4}$$

它的各阶主子式为

$$\Delta_1 = a_1, \quad \Delta_2 = \begin{vmatrix} a_1 & 1 \\ a_3 & a_2 \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix}, \quad \cdots,$$

$$\Delta_n = a_n \Delta_{n-1}.$$

定理 2.2.3 实系数代数方程 (2.2.3) 的所有根都为负实部的充要条件是 (2.2.4) 的一切主子式 $\Delta_i > 0$ $(i = 1, 2, \cdots, n)$.

对于具体的二次、三次、四次方程有如下判据:

(1) 实系数二次方程

$$\lambda^2 + a_1\lambda + a_2 = 0$$

的一切根均具负实部的充要条件是 $a_1 > 0, a_2 > 0$.

(2) 实系数三次方程

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0$$

的一切根均具负实部的充要条件是 $a_1 > 0, a_3 > 0, a_1a_2 - a_3 > 0$.

(3) 实系数四次方程

$$\lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4 = 0$$

的一切根都具有负实部的充要条件是

$$a_1 > 0, \quad a_1a_2 - a_3 > 0, \quad a_3(a_1a_2 - a_3) - a_1^2a_4 > 0, \quad a_4 > 0.$$

第二条途径: 直接对非线性系统 (2.2.1) 构造辅助的 Lyapunov 函数, 利用 Lyapunov 基本定理判别 (2.2.1) 的零解的稳定性. 此法称为 **Lyapunov 直接法**.

研究 (2.2.1) 的零解的稳定性, 就是讨论相空间中原点附近的轨线是怎样随着时间变化而变化的, 原点附近的轨线是趋向于原点还是远离原点. 但问题是我们仅仅知道轨线的方向场, 即 (2.2.1) 的右端函数的值的变化趋势, 如果构作一系列包围原点的闭曲面, 例如一族包围原点的同心球, 在球面上每个点处 (2.2.1) 的方向场是知道的, 于是只需看球面的法向与 (2.2.1) 的方向场的方向的夹角, 就可以确定原点的稳定性.

构造一族闭曲面 $V(x) = C$, 作为对于一族包围原点的同心球的自然推广. 设外面的 C 值比里面的大, 又 $V(0) = 0$. 将相轨线 $x(t)$ 代入函数 $V(x)$ 得到 t 的函数 $V(x(t))$, 它反映了在时刻 t 相轨线 $x(t)$ 经过了哪一个闭曲面. 如果 $V(x(t))$ 是 t 的单调减少函数且趋于零, 则表示相轨线 $x(t)$ 由外向内穿过这一族闭曲面趋向于平衡点 $x = 0$, 即 (2.2.1) 的零解是渐近稳定的. 如果 $V(x(t))$ 的数值

不增加，那么相轨线 x 就不会越出初始的闭曲面 $V(x)=V(x(t_0))$，因此，初始时刻与原点 $x=0$ 充分接近的相轨线 $x(t)$ 仍然在原点附近，即 $x=0$ 是稳定的．如果 $V(x(t))$ 的数值既不增加又不减少，那么相轨线 $x(t)$ 就落在某闭曲面上，即零解附近的轨线落在离 $x=0$ 很接近的闭曲面上，但又不能任意接近原点，所以零解是稳定的但不是渐近稳定的.

为了研究不稳定性，仍然可以用上述的方法．如果我们仍利用上述的一族曲面，而知道相轨线 $x(t)$ 在 V 函数的数值 $V(x(t))$ 随 t 不断增长，则 $x(t)$ 远离原点，零解不稳定．由于不稳定性只要求在原点的任意邻域内存在远离原点的相轨线，而并不要求一切相轨线都远离原点，因此，用 V 函数讨论零解的不稳定性时，只要在原点附近的一部分区域中建立一族曲面，如果在原点附近的任意邻域中存在相轨线离开这个区域，$V(x(t))$ 不断增加，则零解就是不稳定的.

综上所述，可以用一族曲面 $V(x)=C$ 在相空间 R^n 中建立一种广义的尺度，用它来度量相轨线 $x(t)$ 与原点之间的位置关系，确定相轨线的走向，在本质上它与 $\|x(t)\|$ 作用是相似的，只不过前者不必知道具体的 $x(t)$，后者必须知道具体的 $x(t)$.

首先定义 Lyapunov 函数．设区域 $C_H=\{x\mid x\in\mathrm{R}^n,\|x\|\leqslant H,H>0\}$，$\mathrm{R}^+=[0,+\infty)$.

定义 2.2.4　设函数 $V:C_H\to\mathrm{R}^+$ 是连续的，若 $V(x)$ 满足

(1) $V(0)=0$;

(2) 对于一切 $x\in C_H$, $x\neq 0$, $V(x)>0$ (<0),

则称 $V(x)$ 是 **定正 (定负) 函数**.

定义 2.2.5　设函数 $V:C_H\to\mathrm{R}^+$ 是连续的，若 $V(x)$ 满足

(1) $V(0)=0$;

(2) 对于一切 $x\in C_H$, $x\neq 0$, $V(x)\geqslant 0$ $(\leqslant 0)$,

则称 $V(x)$ 是 **常正 (常负) 函数**.

定义 2.2.6　设函数 $V:C_H\to\mathrm{R}^+$ 是连续可微的，则 $V(x)$

按方程组 (2.2.1) 对时间 t 的全导数定义为

$$\left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{(2.2.1)} \triangleq \sum_{k=1}^{n} \frac{\partial V(x)}{\partial x_k} f_k(x) = (\mathrm{grad}\ V, f(x)), \tag{2.2.5}$$

其中 (a,b) 表示 a 与 b 的内积.

下面叙述 Lyapunov 直接方法的基本定理.

定理 2.2.4 对于系统 (2.2.1), 如果存在一个定正函数 $V(x)$, 使得它按 (2.2.1) 对时间 t 的全导数 $\frac{\mathrm{d}V}{\mathrm{d}t}$ 是定负的, 则 (2.2.1) 的零解 $x=0$ 是渐近稳定的.

定理 2.2.5 对于系统 (2.2.1), 如果存在一个定正函数 $V(x)$, 使得它按 (2.2.1) 对时间 t 的全导数 $\frac{\mathrm{d}V}{\mathrm{d}t}$ 是常负的, 则 (2.2.1) 的零解 $x=0$ 是稳定的.

定理 2.2.6 对于系统 (2.2.1), 如果存在函数 $V(x)$ 满足:

(1) 对于一切 $\delta>0$, 原点的 δ 邻域 $O_\delta=\{x \mid \|x\|<\delta\}$ 中总有区域 $\{V>0\}$ 中的点, 即 $\{V>0\}\cap O_\delta\neq\emptyset$, 其中 $\{V>0\}=\{x \mid V(x)>0, x\in C_H\}$, 又 $V(0)=0$;

(2) $V(x)$ 按 (2.2.1) 关于 t 的全导数 $\frac{\mathrm{d}V}{\mathrm{d}t}$ 定正, 则 (2.2.1) 的零解 $x=0$ 是不稳定的.

注意 $V(x)$ 按 (2.2.1) 关于 t 的全导数

$$\left.\frac{\mathrm{d}V(x(t))}{\mathrm{d}t}\right|_{(2.2.1)} = (\mathrm{grad}\ V, f(x(t))),$$

对于曲面 $V(x)=C$, grad V 表示曲面的外法向量, 而 $f(x(t))=\frac{\mathrm{d}x(t)}{\mathrm{d}t}$ 表示相轨线 $x(t)$ 在时刻 t 时所在点处的切向量. 当 $\frac{\mathrm{d}V(x(t))}{\mathrm{d}t}<0$ 时, 即外法向量 grad V 与切向量 $f(x(t))$ 的夹角是钝角, 因此相轨线 $x(t)$ 从包围原点的曲面 $V(x)=C$ 的外部进入内部. 又根据定正函数 $V(x)$ 的几何意义, 这簇闭曲面 $V(x)=C$ 一个包围

一个，随着 $C \to 0$, $V(x) = C$ 收缩于原点，因此轨线 $x(t) \to 0$ $(t \to +\infty)$, 即零解为渐近稳定的.

若 $\dfrac{\mathrm{d}V(x(t))}{\mathrm{d}t} = 0$, 则 grad V 与 $f(x(t))$ 垂直，此时相轨线 $x(t)$ 可能一直停留在曲面 $V(x) = C$ 上. 但是由于 $V(x)$ 定正, 当 $C \to 0$, $V(x) = C$ 收缩于原点，所以保证零解是稳定的.

对于零解不稳定情形， $\dfrac{\mathrm{d}V(x(t))}{\mathrm{d}t} > 0$, 表明 grad V 与 $f(x(t))$ 的夹角为锐角, 即相轨线从包围原点的曲面 $V(x) = C$ 的内部指向外部，且在原点的任意邻域内存在 $V > 0$ 的点，故存在某条相轨线 $x(t)$ 不断由内向外，远离原点.

这三种情形的几何图形见图 2.2.1.

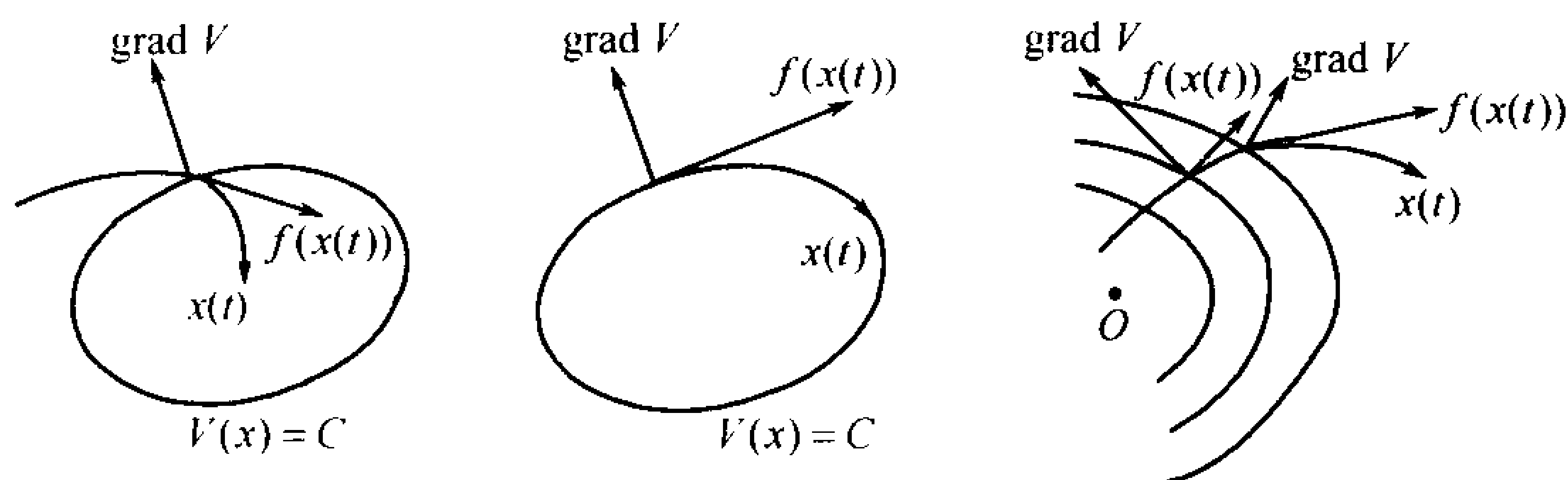


图 2.2.1　grad V 与 $f(x(t))$ 的位置关系

由三个基本定理可知辅助的 V 函数的符号特征有一定的要求. 对渐近稳定、稳定的定理要求 V 正定，对不稳定定理要求原点任意邻域中均有 $V > 0$ 的点，这也是应用时的局限性. 下面一节将介绍变号的 V 函数.

作 Lyapunov 函数的方法有以下几种:

(1) 作二次型的 V 函数，适用于 (2.2.1) 为线性的情形

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Ax. \tag{2.2.6}$$

如果 A 的特征值均具负实部, 那么对于每一个正定对称阵 C $(C > 0)$ 矩阵方程

$$A^{\mathrm{T}}B + BA = -C \tag{2.2.7}$$

存在正定对称阵 B 的解. 因此可以构造二次型的 V 函数

$$V(x) = x^{\mathrm{T}} B x. \tag{2.2.8}$$

易证明

$$\left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{(2.2.6)} = -x^{\mathrm{T}} C x < 0.$$

由基本定理知, (2.2.6) 的零解为渐近稳定的, 并由此可以证明以 (2.2.6) 为线性化的原非线性方程组的零解是渐近稳定的, 所构造的 V 函数仍为 (2.2.8) 的形式.

(2) 能量法. 把系统的能量函数作为 Lyapunov 函数, 根据系统的总的机械能守恒定律, 即 $V(x) \equiv$ 常数, 有 $\dfrac{\mathrm{d}V}{\mathrm{d}t} \equiv 0$, 从而得到零解是稳定的. 如果整个系统存在阻尼, 则在运动过程中, 总的能量是耗散的, 机械能将转化为热能与其它的能量, 这时的 $\dfrac{\mathrm{d}V}{\mathrm{d}t}$ 反映出能量的变化率, 由此容易得出零解是渐近稳定的充分条件.

下面以单摆运动方程为例说明这种方法.

例 2.2.1 先讨论无阻尼的单摆运动方程

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \frac{g}{l}\sin x = 0 \tag{2.2.9}$$

的零解稳定性. 置 $\dfrac{\mathrm{d}x}{\mathrm{d}t} = y$, 得到

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = y, \\[2ex] \dfrac{\mathrm{d}y}{\mathrm{d}t} = -\dfrac{g}{l}\sin x, \end{cases}$$

系统的总的机械能为

$$E = \frac{1}{2}m(ly)^2 + mgl(1 - \cos x) = \text{常数}. \tag{2.2.10}$$

置 $V(x,y)=\dfrac{1}{2}m(ly)^2+mgl(1-\cos x)$, 它是定正函数. 而 V 按照方程组对 t 的全导数

$$\frac{\mathrm{d}V}{\mathrm{d}t}=(mgl\sin x)y+ml^2y\left(-\frac{g}{l}\sin x\right)\equiv 0.$$

由基本定理得零解是稳定的. 由 $V(x(t),y(t))\equiv V(x_0,y_0)$, 所以零解不是渐近稳定的.

其次考虑有阻尼的摆的运动方程

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2}+c\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{g}{l}\sin x=0, \tag{2.2.11}$$

其中阻尼系数 $c>0$. 我们知道, 有阻尼的摆一旦偏离平衡位置 $x=0$, $\dfrac{\mathrm{d}x}{\mathrm{d}t}=0$, 当经过很长一段时间后, 摆将趋于平衡位置, 因此实际系统是渐近稳定的. 此时微分方程化为如下的一阶方程组

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=y,\\[2ex]\dfrac{\mathrm{d}y}{\mathrm{d}t}=-cy-\dfrac{g}{l}\sin x.\end{cases}$$

若用能量法, 仍构造 V 函数为

$$V(x,y)=\frac{1}{2}m(ly)^2+mgl(1-\cos x), \tag{2.2.12}$$

则 $\dfrac{\mathrm{d}V}{\mathrm{d}t}=-mcl^2y^2$, 它仅是常负函数, 故只能判别其零解是稳定的, 而不能判别其零解为渐近稳定的. 为此要引入判别零解渐近稳定的 Krasovskii–Barbasin 定理.

先引进正半轨的概念: 非平衡点的解轨线 $x(t;t_0,x_0)$ 中 $t\geqslant t_0$ 的部分称为 **正半轨**.

定理 2.2.7 (Krasovskii–Barbasin 定理) 对于自治系统 (2.2.1), 如果存在定正函数 $V(x)$, 使得 $V(x)$ 沿 (2.2.1) 关于 t 的全

导数 $\dfrac{\mathrm{d}V}{\mathrm{d}t}$ 是常负的，而集合 $\left\{x \mid x \in C_H, \dfrac{\mathrm{d}V}{\mathrm{d}t}=0\right\}$ 不包含 (2.2.1) 的非零正半轨，则 (2.2.1) 的零解是渐近稳定的.

回到例 2.2.1 有阻尼的情形，此时全导数 $\dfrac{\mathrm{d}V}{\mathrm{d}t}$ 为常负的，但是使得 $\dfrac{\mathrm{d}V}{\mathrm{d}t}=0$ 的 x 集合为 $\{(x,0)\}$, 显然除 (0,0) 外它不是 (2.2.1) 的正半轨，即满足定理 2.2.7, 故知 (2.2.1) 的零解是渐近稳定的.

2. 时滞型的泛函微分方程模型与 Lyapunov 泛函方法

考虑如下的自治时滞型泛函微分方程 (RFDE)

$$\dot{x}(t)=f(x_t), \tag{2.2.13}$$

其中 $f: C \to \mathrm{R}^n$ 为连续的，且 $f(t,0)\equiv 0$.

现定义 Lyapunov 泛函如下：设泛函 $V: C \to \mathrm{R}^+$ 连续，且它沿 (2.2.13) 的解的右上导数定义为

$$\dot{V}(\phi) \triangleq \frac{\mathrm{d}V}{\mathrm{d}t}(\phi)\bigg|_{(2.2.13)} = \overline{\lim_{h\to 0^+}}\frac{1}{h}\{V(x_{t+h}(\phi))-V(\phi)\}. \tag{2.2.14}$$

又设 $u,v,w: \mathrm{R}^+ \to \mathrm{R}^+$ 为连续非减函数，当 $s>0$ 时， $u(s)>0$, $v(s)>0$, $w(s)>0$, $u(0)=v(0)=w(0)=0$.

下面介绍几个判别 (2.2.13) 零解稳定与渐近稳定的 Lyapunov 泛函方法的几个基本定理. 为了避免一些复杂的专业性的术语与记号，这里介绍的是已经简化了的结果.

定理 2.2.8 若存在上述的函数 u, v 以及 Lyapunov 泛函 $V(\phi)$ 满足下列条件:

(1) 对任意 $\phi \in C$, 有

$$u(|\phi(0)|) \leqslant V(\phi) \leqslant v(\|\phi\|);$$

(2) 对任意 $t_0 \geqslant 0$ 及 $\phi \in C$, 有

$$\dot{V}(x_t(t_0,\phi)) \leqslant 0, \quad t \geqslant t_0,$$

则方程 (2.2.13) 的零解为稳定的.

注 2.2.1 实际上这个定理的结论可以改为零解是一致稳定的.

定理 2.2.9 如果存在上述的函数 u, v 以及 Lyapunov 函数 $V(x)$ 满足如下条件:

(1) $u(|x|) \leqslant V(x) \leqslant v(|x|)$, $x \in \mathrm{R}^n$;

(2) 对任何 $t_0 \geqslant 0$ 及 $\varphi \in C$, 当 $t-\tau \leqslant \xi \leqslant t$, $V(x(t_0,\varphi)(t)) \geqslant V(x(t_0,\varphi)(\xi))$ 时, 有

$$\dot{V}(x(t_0,\varphi)(t)) \leqslant 0,$$

则方程 (2.2.13) 的零解稳定.

注 2.2.2 定理 2.2.9 的结果可以加强为 (2.2.13) 零解一致稳定. 该定理用的方法是 Lyapunov 函数的方法而不是 Lyapunov 泛函的方法, 人们常称之为 Ragumikhin 型定理.

举一个例子说明这两个定理的应用.

例 2.2.2 考虑纯量方程

$$\dot{x}(t) = -ax(t) + bx(t-\tau) \tag{2.2.15}$$

的解的稳定性, 其中 $a > b > 0$.

方法 1 使用 Lyapunov 泛函. 取

$$V(\phi) = \frac{1}{2}\phi^2(0) + \mu\int_{-\tau}^{0}\phi^2(\theta)\mathrm{d}\theta,$$

$\mu > 0$ 为待定. 于是

$$\begin{aligned}\frac{\mathrm{d}V(x_t)}{\mathrm{d}t} &= x(t)\frac{\mathrm{d}x(t)}{\mathrm{d}t} + \mu\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{-\tau}^{0}x^2(t+\theta)\mathrm{d}\theta\right] \\ &= x(t)[-ax(t)+bx(t-\tau)] + \mu\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{t-\tau}^{t}x^2(s)\mathrm{d}s\right]\end{aligned}$$

$$
\begin{aligned}
&= x(t)[-ax(t) + bx(t-\tau)] + \mu[x^2(t) - x^2(t-\tau)] \\
&= -(a-\mu)x^2(t) + bx(t)x(t-\tau) - \mu x^2(t-\tau) \\
&= -(a-\mu)\left[x(t) - \frac{b}{2(a-\mu)}x(t-\tau)\right]^2 \\
&\quad + \frac{b^2 - 4\mu(a-\mu)}{4(a-\mu)}x^2(t-\tau).
\end{aligned}
$$

取 $\mu = \dfrac{a}{2}$, 有

$$
\dot{V}(x_t) = -\frac{a}{2}\left[x(t) - \frac{b}{a}x(t-\tau)\right]^2 + \frac{b^2 - a^2}{2a}x^2(t-\tau) \leqslant 0.
$$

若取 $u(s) = \dfrac{s^2}{2}$, $v(s) = \left(\dfrac{1}{2} + \dfrac{a\tau}{2}\right)s^2$, 则易见定理 2.2.8 的两个条件满足，从而得证 (2.2.15) 的零解是稳定的.

方法 2 使用 Lyapunov 函数. 取

$$
V(x) = \frac{x^2}{2},
$$

于是

$$
\begin{aligned}
\dot{V}(x(t)) &= -ax^2(t) + bx(t)x(t-\tau) \\
&\leqslant -ax^2(t) + b|x(t)||x(t-\tau)| \\
&\leqslant -ax^2(t) + \frac{b}{2}x^2(t) + \frac{b}{2}x^2(t-\tau).
\end{aligned}
$$

当 $V(x(t)) \geqslant V(x(\xi))$, $t-\tau \leqslant \xi \leqslant t$ 时，有 $V(x(t)) \geqslant V(x(t-\tau))$, 即有 $x^2(t) \geqslant x^2(t-\tau)$, 故

$$
\begin{aligned}
\dot{V}(x(t)) &\leqslant -ax^2(t) + \frac{b}{2}x^2(t) + \frac{b}{2}x^2(t) \\
&= -(a-b)x^2(t) \leqslant 0.
\end{aligned}
$$

如果取 $u(s) = \dfrac{s^2}{3}$, $v(s) = s^2$, 那么易知定理 2.2.9 的两个条件满

足, 所以 (2.2.15) 的零解是稳定的.

定理 2.2.10 若存在上述的函数 u, v, w 以及 Lyapunov 泛函 $V(\phi)$ 满足下列条件:

(1) 对任意 $\phi \in C_H = \{\phi \in C : \|\phi\| \leqslant H\}$, 有

$$u(|\phi(0)|) \leqslant V(\phi) \leqslant v(\|\phi\|);$$

(2) $\dot{V}(x_t) \leqslant -w(|x(t)|)$;

(3) $f(\phi)$ 在 C_H 有界,

则 (2.2.13) 的零解为渐近稳定的.

注 2.2.3 定理 2.2.10 的结果可以加强为 (2.2.13) 的零解为一致渐近稳定的. 条件 (3) 可以减弱为 $[\phi(0)]^{\mathrm{T}}$ 在 C_H 上有上界或有下界.

定理 2.2.11 若存在上述的函数 u, v, w 以及 Lyapunov 函数 $V(x)$ 满足下列条件:

(1) $u(|x|) \leqslant V(x) \leqslant v(|x|)$, $|x| \leqslant H$;

(2) 有某函数 $P(z)$, 使得 $P(V(x(t)) > V(x(\xi))$, $t-\tau \leqslant \xi \leqslant t$ 时有

$$\dot{V}(x(t)) \leqslant -W(|x(t)|),$$

则 (2.2.13) 的零解为渐近稳定的.

注 2.2.4 定理 2.2.11 的结果可以加强为 (2.2.13) 的零解为一致渐近稳定的. 本定理是 Ragumikhin 型的一个典型定理, 用的方法是 Lyapunov 函数方法, 并不涉及到 Lyapunov 泛函.

下面举例说明这两个渐近稳定的判别定理和应用.

例 2.2.3 考虑 n 维的时滞系统

$$\dot{x}(t) = Ax(t) + Bx(t-\tau), \quad \tau > 0, \tag{2.2.16}$$

其中 $x(t)$ 是 n 维向量值函数, A 和 B 是 $n \times n$ 常数矩阵. 假定 A 的所有特征值均具有负实部.

现取 V 泛函为

$$V(\phi) = \phi^{\mathrm{T}}(0)C\phi(0) + \int_{-\tau}^{0} \phi^{\mathrm{T}}(\theta)E\phi(\theta)\mathrm{d}\theta, \tag{2.2.17}$$

其中 T 表示转置， E 和 C 皆为 $n \times n$ 正定对称矩阵，且使得

$$\begin{aligned} &A^{\mathrm{T}}C + CA = -D < 0 \quad (D \text{ 为任一给定的正定矩阵}), \\ &\dot{V}(x_t)|_{(2.2.16)} = -x^{\mathrm{T}}(t)(D-E)x(t) + 2x^{\mathrm{T}}(t)CBx(t-\tau) \\ &\qquad\qquad\quad -x^{\mathrm{T}}(t-\tau)Ex(t-\tau). \end{aligned}$$

将上式右端看作 $x(t)$ 和 $x(t-\tau)$ 的二次型，假定 $E-D$ 为负定，记之为 $E-D<0$, 即 $D-E>0$. 又设

$$\begin{aligned} x^{\mathrm{T}}(t)(D-E)x(t) &\geqslant \lambda|x(t)|^2 \quad (\lambda > 0), \\ x^{\mathrm{T}}(t)Ex(t) &\geqslant \mu|x(t)|^2 \quad (\mu > 0). \end{aligned}$$

若 (2.2.16) 中的 B 满足 $\|B\| < \dfrac{\sqrt{\lambda\mu}}{\|C\|}$, 即有 $\|CB\|^2 < \lambda\mu$, 则存在常数 $k > 0$, 使得

$$\dot{V}(x_t) \leqslant -k(|x(t)|^2 + |x(t-\tau)|^2) \leqslant -k|x(t)|^2.$$

定义 $w(s) = ks^2$, 则有

$$\dot{V}(x_t) \leqslant -w(|x(t)|),$$

从而满足定理 2.2.10 的 3 个条件，所以 (2.2.16) 的零解是渐近稳定的.

2.2.2 极限集及 Lasalle 不变性原理

对于差分方程、常微分方程以及泛函微分方程 Lasalle 不变性原理有其不同的形式，我们将分别给予介绍.

1. 差分方程情形 Lasalle 不变性原理的应用

首先介绍一些基本概念：极限集、不变集，并给出差分方程情形的 Lyapunov 函数及其全导数的定义，然后给出差分方程的不变性原理，最后通过实例说明如何具体应用.

定义 2.2.7 对差分方程

$$x_{n+1} = T(x_n), \quad n \in I^+ = \{0, 1, 2, \cdots\}, \tag{2.2.18}$$

其中 $T: \mathrm{R}^m \to \mathrm{R}^m$, y 是 $T^n x$ 的 **极限点**, 是指若存在整数 n_i, 使得 $T^{n_i} x \to y$ 与 $n_i \to \infty$, 当 $i \to \infty$ 时. $T^n x$ 的 **极限集** $\Omega(x)$ 为 $T^n x$ 的所有极限点的集合.

实际上有 $\Omega(x) = \bigcap\limits_{j=0}^{\infty} \overline{\bigcup\limits_{n=j}^{\infty} T^n x}$.

上面用到的记号 $\overline{S}$, 是指某一集合 S 的闭包.

定义 2.2.8 对 (2.2.18) 的映射 T, 集合 T 是 **正不变 (负不变)** 的定义为 $T(H) \subset H$ $(H \subset T(H))$. 若 $T(H) = H$, 则称 H 是 **不变** 的.

极限集的基本性质可用下面的定理来表达.

定理 2.2.12 每个极限集 $\Omega(x)$ 均是闭的和正不变的. 如果 $T^n x$ 对 $n \in I^+$ 是有界的, 那么 $\Omega(x)$ 是非空、紧致、不变地连通的, 又是 $T^n x$ 当 $n \to \infty$ 时趋向的最小闭集.

置 $V: \mathrm{R}^m \to \mathrm{R}$, 对 (2.2.18) 的映射 T 定义

$$\dot{V}(x) = V(T(x)) - V(x).$$

如果 (2.2.18) 的解为 x_n, 那么

$$\dot{V}(x_n) = V(x_{n+1}) - V(x_n),$$

而 $\dot{V}(x) \leqslant 0$ 意味着 V 沿 (2.2.18) 的解为非增的.

定义 2.2.9 设 $G \subset \mathrm{R}^m$, 称 V 是 (2.2.18) 在 G 上的 **Lyapunov 函数** 指 V 是连续的且对所有 $x \in G$ 有 $\dot{V}(x) \leqslant 0$.

引进如下 3 个集合

$$E=\{x:\dot{V}(x)=0,x\in\overline{G}\},\quad \overline{G}\text{ 为 }G\text{ 的闭包,}$$
$$M=E\text{ 中的最大不变集,}$$
$$V^{-1}(c)=\{x:V(x)=c,x\in\mathrm{R}^m\}.$$

定理 2.2.13 (差分方程映射的不变性原理) 如果 V 是 (2.2.18) 在 G 上的 Lyapunov 函数, 而且 (2.2.18) 的解 x_n 对任意 $n\geqslant 0$ 均有界且在 G 中, 那么存在数 c, 使得当 $n\to\infty$ 时, $x_n\to M\cap V^{-1}(c)$.

在应用该定理时, 要求出 $M\cap V^{-1}(c)$. 该集合可能是 (2.2.18) 的平衡点集, 也可能是 (2.2.18) 的周期点集极限环.

例 2.2.4 考虑如下的二维系统

$$\begin{cases} x_{n+1}=\dfrac{ay_n}{1+x_n^2}, \\ y_{n+1}=\dfrac{bx_n}{1+y_n^2}, \end{cases} \tag{2.2.19}$$

其中 a,b 为常数.

取 $V(x,y)=x^2+y^2$, 则

$$\begin{aligned}\dot{V}(x,y)&=V(T(x,y))-V(x,y)\\&=\left[\frac{ay}{1+x^2}\right]^2+\left[\frac{bx}{1+y^2}\right]^2-x^2-y^2\\&=\left[\frac{b^2}{(1+y^2)^2}-1\right]x^2+\left[\frac{a^2}{(1+x^2)^2}-1\right]y^2.\end{aligned}$$

(1) 当 $a^2<1,b^2<1$ 时, 有

$$\dot{V}(x,y)\leqslant(b^2-1)x^2+(a^2-1)y^2<0.$$

易知 V 是 (2.2.19) 在 R^2 上的 Lyapunov 函数, 这里 $M=E=\{(0,0)\}$, 也就有 $M\cap V^{-1}(c)=\{(0,0)\}$. 显然 (2.2.19) 的每个解是

有界的, 由定理 2.2.13 知, 当 $n \to \infty$ 时, (2.2.19) 的解 $(x_n, y_n) \to (0,0)$, 即原点是整体渐近稳定的或原点是整体吸引子.

(2) 当 $a^2 < 1, b^2 = 1$ 时, V 是 (2.2.19) 在 R^2 上的 Lyapunov 函数, 但 $\dot{V} \leqslant (a^2-1)y^2 \leqslant 0$, 这里 $E = \{(x,0)\}$, 即 E 为 x 轴. 又因为 $T(x,0) = (0, bx)$, 所以 E 中的最大不变集 $M = \{(0,0)\}$, 即知 $M \cap V^{-1}(c) = \{(0,0)\}$. 由定理 2.2.13 知, 当 $n \to \infty$ 时, (2.2.19) 的解 $(x_n, y_n) \to (0,0)$, 即原点是整体渐近稳定的.

(3) 当 $a^2 = b^2 = 1$ 时, (2.2.19) 的解仍是有界的. 易证 $E = M = \{(x,0),(0,y)\}$, 即 E 与 M 为两个坐标轴. 又 $M \cap V^{-1}(c^2) = \{(c,0),(-c,0),(0,c),(0,-c)\}$, 其中 $c > 0$. 分下面两种情形讨论:

(i) $ab = 1$ 的情形. 注意 $T(\pm c, 0) = (0, \pm bc)$,

$$T^2(\pm c, 0) = T(0, \pm bc) = (\pm abc, 0) = (\pm c, 0).$$

又 $T(0, \pm c) = (\pm ac, 0)$, $T^2(0, \pm c) = T(\pm ac, 0) = (0, \pm c)$.

因为极限集是不变地连通的, 所以 (2.2.19) 的每个解趋向于原点 $(0,0)$ 或周期 2 点 $(\pm c, 0)$ 或 $(0, \pm c)$, 这里 $c > 0$.

(ii) $ab = -1$ 的情形. 注意 $T(c,0) = (0, bc)$,

$$\begin{aligned}
T^2(c,0) &= T(0,bc) = (abc,0) = (-c,0), \\
T^3(c,0) &= T(-c,0) = (0,-bc), \\
T^4(c,0) &= T(0,-bc) = (-abc,0) = (c,0).
\end{aligned}$$

当 $c > 0$, 即 $(c,0)$ 或 $(-c,0)$ 均为周期 4 点, 因为极限集是不变地连通的, 所以 (2.2.19) 的每个解趋向于原点 $(0,0)$ 或周期 4 点 $(c,0)$ 或 $(-c,0)$ 或 $(0,c)$ 或 $(0,-c)$, 这里 $c > 0$.

(4) 当 $a^2 > 1$, $b^2 > 1$ 时, 置 $B_\delta = \{(x,y) : x^2 + y^2 < \delta^2\}$. 对 $(x,y) \in B_\delta$ 和 δ 充分小, 有

$$\dot{V} \geqslant \left[\frac{b^2}{(1+\delta^2)^2} - 1\right] x^2 + \left[\frac{a^2}{(1+\delta^2)^2} - 1\right] y^2 \geqslant 0,$$

且 $-V$ 是 (2.2.19) 在 B_δ 上的 Lyapunov 函数. $E = M = \{(0,0)\}$. 从 B_δ 中异于原点的点出发, V 的值是增加的, 又 V 为 (x,y) 离原

点的距离，由 $T(x,y)=(0,0)$ 推出 $x=y=0$，知原点是平衡点.又 $M\cap V^{-1}(c)=\emptyset$，由定理 2.2.13 知 (2.2.19) 的解趋向于空集，即没有 (2.2.19) 的非零解趋向于原点，而任何非零解将离开 B_δ.

从该例可知，Lasalle 不变性原理可以用于判别系统的解的稳定性与不稳定性，既可用于平衡点不动点吸引子与排斥子的判别，也可用于极限环周期点吸引子与排斥子的判别. 该原理可以整体判别平衡点集的吸引性，克服了 Lyapunov 直接方法中某个平衡点解的稳定性的局部判别的限制. 实际上，用该原理时不必要求构造的 Lyapunov 函数为定号，因此 Lyapunov 基本定理中对函数定号的限制可以去掉，而许多系统包括神经网络系统的广义能量函数恰恰不是定号的，但按 Lasalle 不变性原理照样可以拿来用于吸引稳定性的判别.

由 Lasalle 不变性原理可以引出几个实际判别中有用的定理. 这里首先给出一些概念的定义.

定义 2.2.10 (1) 集合 H 称为是 **稳定** 的指若对给定的 H 的邻域 U，即包含 $\overline{H}$ (H 的闭包) 的开集，存在 H 的邻域 W，使得对所有 $n\in J_+$，有 $T^n(W)\subset U$;

(2) 集合 H 是 **吸引子** 指存在 $\overline{H}$ 的邻域 U，使得当 $x\in U$ 时有 $\lim\limits_{n\to\infty}T^nx\in\overline{H}$;

(3) H 称为是 **渐近稳定** 的指如果 H 既是稳定的又是吸引子;

(4) 如果 H 是渐近稳定的，又对所有 $x\in \mathrm{R}^m$ 有 $\lim\limits_{n\to+\infty}T^nx\in\overline{H}$ 成立，那么称 H 是 **整体渐近稳定** 的;

(5) 如果 H 既不是稳定的又不是吸引子，那么 H 是 **强不稳定** 的.

定理 2.2.14 设 G 是有界开的正不变集. 如果

(1) V 是 (2.2.18) 在 G 上的 Lyapunov 函数,

(2) $M\subset G$,

那么 M 是吸引子，而且 $\overline{G}$ (G 的闭包) $\subset R(M)$ (M 的吸引域).

如果除上述两个条件外还满足

(3) V 在 M 上是常数,

那么 M 是渐近稳定的 (相对 G 的整体渐近稳定的).

注 2.2.5 如果 M 是单点集或者是仅带有限个元素的不变的连通的集合, 那么定理 2.2.14 的条件 (3) 自动满足, 即我们得到的判别 M 是渐近稳定的充分条件既不要求 V 定正也不要求 $-\dot{V}$ 定正.

定理 2.2.15 设 G 是有界开集且关于 (2.2.18) 是正不变的. 如果 V 是 (2.2.18) 在 G 上的 Lyapunov 函数且对 $x \in G, x \neq 0$ 成立 $\dot{V}(x) < 0$. 又置 E_0 是 G 的边界上使 $\dot{V} = 0$ 的点集, M_0 是 E_0 中的最大不变集, 则从 G 中出发的所有解趋向于 $M = M_0 \cup \{0\}$.

(1) 若 $M = \{0\}$, 则原点是相对于 G 为整体渐近稳定的;

(2) 若 $\{0\} \in G$, 则 G 中出发的每个解当 $n \to \infty$ 时或趋向于 0 或趋向于 M_0. 如果没有解当 $n \to \infty$ 时趋向于 M_0, 则原点是渐近稳定的或者是相对于 G 为整体渐近稳定的.

定理 2.2.16 设 H 是紧致的, 又 G 是包含 H 的开集. 如果

(1) 对 $x \in H$ 成立 $V(x) \leqslant 0$, 对 $x \in G - H$ 成立 $V(x) > 0$;

(2) V 是 (2.2.18) 在 G 上的 Lyapunov 函数,

则 H 是稳定的;

又若除了上述 (1), (2) 外还满足

(3) $E \subset H$,

则 H 是渐近稳定.

例 2.2.5 考虑如下模型

$$\begin{cases} x_1(n+1) = a_1 x_2(n)(1 - x_1(n)) + (1 - b_1)x_1(n), \\ x_2(n+1) = a_2 x_1(n)(1 - x_2(n)) + (1 - b_2)x_2(n), \end{cases} \tag{2.2.20}$$

其中 $0 < a_i < 1, 0 < b_i < 1$ $(i = 1, 2)$. 改记 (2.2.20) 为如下形式

$$x' = Tx = A(x)x, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{2.2.21}$$

$$A(x)=\begin{pmatrix} 1-b_1 & a_1(1-x_1) \\ a_2(1-x_2) & 1-b_2 \end{pmatrix}. \tag{2.2.22}$$

置 $G=\{x \mid 0<x_i<1, i=1,2\}$, 易见 G 和 $\overline{G}$ 是正不变集, 实际上有 $T\{\overline{G}-\{0\}\}\subset G$. 把 $A(x)$ 作如下分解:

$$\begin{aligned} A(x)&=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}+\begin{pmatrix} -b_1 & a_1 \\ a_2 & -b_2 \end{pmatrix}+\begin{pmatrix} 0 & -a_1x_1 \\ -a_2x_2 & 0 \end{pmatrix} \\ &=I+A_0+A_1(x). \end{aligned}$$

假设 $(1-b_1)(1-b_2)-a_1a_2 \geqslant 0$, 易知 $A(x)\geqslant 0$, $I+A_0\geqslant 0$, $A_1(x)\leqslant 0$ $(x\in G)$, $A_1(0)=0$.

记 $r(P)$ 为 P 矩阵的特征根的谱半径, 分下面两种情况讨论:

(1) $r(I+A_0)\leqslant 1$ 的情形. 因为 $I+A_0\geqslant 0$, 所以 $-A_0$ 的主子式均为非负的, 即

$$\det A_0=b_1b_2-a_1a_2\geqslant 0.$$

取

$$V(x)=c^{\mathrm{T}}x=\begin{pmatrix} b_2 & a_1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=b_2x_1+a_1x_2,$$

有

$$\begin{aligned} \dot{V}(x)&=c^{\mathrm{T}}(A_0+A_1(x))x \\ &=(a_1a_2-b_1b_2)x_1-a_1(a_2+b_2)x_1x_2\leqslant 0, \quad x\in\overline{G}, \end{aligned}$$

且 $\dot{V}(x)<0$, $x\in G$,

$$\begin{aligned} &E_0=\{(0,x_2), 0\leqslant x_2\leqslant 1\}, \\ &M_0=\{(0,0)\}, \\ &M=M_0\cup\{(0,0)\}=\{(0,0)\}. \end{aligned}$$

由定理 2.2.15 知原点是相对于 $\overline{G}$ 整体渐近稳定的.

(2) $r(I+A_0)>1$ 的情形. 对 (2.2.21), (2.2.22) 的一次近似系统

$$x' = (I + A_0)x \tag{2.2.23}$$

原点是不稳定的. 注意此时在 G 中原点的附近 $V(x)$ 与 $\dot{V}(x)$ 均是正的, 从而没有从 $\overline{G}-\{0\}$ 出发的解当 $n\to\infty$ 时能趋向于原点. 实际上此时所有这种解趋向于 G 中的不变集, 即 G 中的唯一的平衡点

$$x_1^0 = \frac{1-\gamma_1\gamma_2}{1+\gamma_1}, \quad x_2^0 = \frac{1-\gamma_1\gamma_2}{1+\gamma_2},$$

其中 $\gamma_i = \dfrac{b_i}{a_i}$ $(i=1,2)$. 由于 $r(I+A_0)>1$, 此时 $\det A_0 = b_1b_2 - a_1a_2 < 0$, 即 $\gamma_1\gamma_2 < 1$. 关于平衡点 (x_1^0, x_2^0) 作平移变换 $u = x - x^0 = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $x_0 = \begin{pmatrix} x_1^0 \\ x_2^0 \end{pmatrix}$, 则可得

$$u' = B(x)u, \tag{2.2.24}$$

其中

$$B(x) = \begin{pmatrix} 1-a_1\rho & a_1(1-x_1) \\ a_2(1-x_2) & 1-a_2\rho^{-1} \end{pmatrix},$$

$$\rho = \frac{1+\gamma_1}{1+\gamma_2}.$$

注意

$$\begin{pmatrix} a_2 & a_1\rho \end{pmatrix}\begin{pmatrix} -a_1\rho & a_1 \\ a_2 & -a_2\rho^{-1} \end{pmatrix} = (0,0),$$

取

$$V(u_1,u_2) = a_2|u_1| + a_1\rho|u_2|.$$

如果 $1-a_1\rho \geqslant 0$, $1-a_2\rho^{-1}\geqslant 0$, 那么

$$\dot{V}(u_1,u_2) \leqslant -a_1a_2(x_1|u_2| + \rho x_2|u_1|) \leqslant 0, \quad x\in G.$$

实际上若 $a_i + b_i \leqslant 1$ $(i = 1, 2)$, 则有 $1 - a_1\rho \geqslant 0$, $1 - a_2\rho^{-1} \geqslant 0$. 此时定理 2.2.15 中的集合 M_0 是原点, 由该定理可知 $\overline{G} - \{0\}$ 中出发的解没有一个能够趋向于 M_0, 即 $u = 0$ (对应 x^0) 相对于 $G - \{0\}$ 是渐近稳定的.

2. 常微分方程情形 Lasalle 不变性原理的应用

对于 n 维自治的常微分方程

$$\dot{x} = f(x), \quad x \in \mathrm{R}^n, \tag{2.2.25}$$

$f : G^* \to \mathrm{R}^n$, G^* 是 R^n 中的一个开集.

首先给出极限集、不变集与 Lyapunov 函数的定义, 然后介绍 (2.2.25) 的 Lasalle 不变性原理的结果, 最后通过例子说明该原理的应用.

假设 f 是连续的, $\pi(t, x^0)$ 是 (2.2.25) 的 G^* 中满足初始条件 $\pi(0, x^0) = x^0$ 的唯一的解.

定义 2.2.11 设 $\phi : (\alpha, \omega) \to G^*$, 其中 $-\infty \leqslant \alpha < 0 < \omega \leqslant +\infty$. 点 p 称为 ϕ 的 **正的 (负的) 极限点**, 是指若存在序列 $t_n \in (\alpha, \omega)$, 使得当 $n \to \infty$ 时, 有 $t_n \to \omega$ $(t_n \to \alpha)$ 和 $\phi(t_n) \to p$. p 的所有的正的 (负的) 极限点的集合 $\Omega(\phi)$ $(A(\phi))$ 称为 ϕ 的 **正的 (负的) 极限集**. 区间 (α, ω) 称为 **最大** 的 (这是与 G^* 相依的), 是指由 $\omega < \infty$ 推出 $\Omega(\phi) \cap G^*$ 是空的和 $-\infty < \alpha$ 推出 $A(\phi) \cap G^*$ 是空的.

定义 2.2.12 (2.2.25) 的满足 $\pi(0, x) = x$, $x \in G^*$ 的解 $\pi(t, x)$ 的最大存在区间 $I(x) = (\alpha(x), \omega(x))$, $-\infty \leqslant \alpha < 0 < \omega \leqslant +\infty$. $\pi(t, x)$ 满足如下的半群及唯一性质: 对所有 $s \in I(x)$ 及 $t \in I(\pi(s, x))$ 有 $t + s \in I(x)$ 及 $\pi(t, \pi(s, x)) = \pi(t + s, x)$. $\pi(t, x)$ 还满足如下的连续依赖性质: 如果 $(t_n, x^n) \in I(x^n) \times G^*$, $(t_n, x^n) \to (t, x) \in I(x) \times G^*$, 则当 $n \to \infty$ 时, 有 $\pi(t_n, x^n) \to \pi(t, x)$.

定义 2.2.13 相对于系统 (2.2.25), 集合 $H \subset \mathrm{R}^n$ 称为 **正的 (负的) 不变** 的, 是指若由 $x \in H \cap G^*$ 可推得对所有 $t \in [0, \omega(x))$

($t \in (\alpha(x), 0]$) 成立 $\pi(t,x) \in H$. H 称为 **弱不变** 的, 是指 H 是正的和负的不变的. 又若对每个 $x \in H \cap G^*$ 有 $I(x) = (-\infty, +\infty)$, 且 H 是正的和负的不变的, 则 H 是不变的.

定义 2.2.14 置 $V: G^* \to \mathrm{R}$, 定义

$$\dot{V}(x) = \liminf_{t \to 0^+} \frac{1}{t}[V(\pi(t,x)) - V(x)], \tag{2.2.26}$$

如果 $\phi(t) = V(\pi(t,x))$, 又 $D_+\phi(t)$ 是右半导数, 那么 $D_+\phi(t) = \dot{V}(\pi(t,x))$. 设 G 是 G^* 的任意子集, 称 V 是 (2.2.25) 在 G 上的 **Lyapunov 函数** 指 V 是连续的, 且对 $x \in G$ 成立 $\dot{V}(x) \leqslant 0$.

定义 2.2.15 对 (2.2.25) 在 G 上的 Lyapunov 函数 V, 定义

$$E = \{x : \dot{V}(x) = 0, x \in \overline{G} \cap G^*\},$$

$M = E$ 中最大不变集,

$M^* = E$ 中最大弱不变集,

$M^+ = E$ 中最大的正不变集.

一般地, $M \subset M^* \subset M^+$ 成立.

定理 2.2.17 (常微分方程的 Lasalle 不变性原理) 设 V 是 (2.2.25) 在 G 上的 Lyapunov 函数, 又 $x(t) = \pi(t, x^0)$ 是对所有 $t \in [0, \omega(x^0))$ 留在 G 中的 (2.2.25) 的解, 则对某个常数 c, $\Omega(x^0) \cap G^* \subset M^* \cap V^{-1}(c)$. 若 $x(t)$ 是预紧致的, 则 $x(t) \to M \cap V^{-1}(c)$ $(t \to \infty)$.

定理 2.2.18 如果 V 是 (2.2.25) 在 G 上的 Lyapunov 函数, 且 $x(t)$ 是对所有 $t \to 0$ 留在 G 中的 (2.2.25) 的预紧致的解, 又 M (或 E) 同 $V^{-1}(c)$ 的交点是对每个 c 均为孤立的, 那么当 $t \to \infty$ 时, $x(t)$ 趋向于 (2.2.25) 的平衡点.

定义 2.2.16 (1) 一个紧致集 $H \subset G^*$ 称为 **稳定** 的, 是指对给定的 H 的邻域 U, 存在 H 的邻域 W, 使得由 $x \in W$ 推得对所有 $t \geqslant 0$ 成立 $\pi(t,x) \in U$;

(2) 一个紧致集 $H \subset G^*$ 称为一个 **吸引子**, 是指存在 H 的邻域 U, 使得由 $x \in U$ 推得当 $t \to \infty$ 时有 $\pi(t,x) \to H$;

(3) 如果对每个 $x \in G^*$ 有 $\pi(t,x) \to H$, 则称 H 为 **整体吸引子**;

(4) 如果 H 既是稳定的又是吸引子, 则 H 是 **渐近稳定** 的;

(5) 如果 H 既是稳定的又是整体吸引子, 那么 H 是 **整体渐近稳定** 的;

(6) 如果 H 不是稳定的又不是吸引子, 那么 H 称为是 **强不稳定** 的;

(7) G^* 中的集合 H 的 **吸引域** $R(H)$ 是使得当 $t \to \infty$ 时成立 $\pi(t,x) \to H$ 的所有 $x \in G^*$ 的集合.

定理 2.2.19 设 G 是 G^* 中的一个正不变开集, 从 G 中出发的 (2.2.25) 的每个解均是有界的而且没有正的极限点是落在 G 的边界上. 如果下面 3 个条件满足:

(1) V 是 G 上的 (2.2.25) 的 Lyapunov 函数;

(2) $M^0 \triangleq \overline{M \cap G} \subset G$;

(3) M^0 是紧致的,

那么 M^0 是一个吸引子, 又 $G \subset R(M^0)$;

如果除了上述 3 个条件外还满足如下第 4 个条件:

(4) V 在 M^0 的边界上是常数,

那么 M^0 是渐近稳定的.

例 2.2.6 考虑系统

$$\begin{cases} \dot{x}_1 = -x_1|x_2|^\alpha, \\ \dot{x}_2 = -x_2|x_1|^\beta, \end{cases} \tag{2.2.27}$$

这里 $\alpha > 0, \beta > 0$. 置 $V = x_1^2 + x_2^2$, 则有

$$\dot{V}(x) = -2(x_1^2|x_2|^\alpha + |x_1|^\beta x_2^2) \leqslant 0,$$

即 V 是 R^2 上的 Lyapunov 函数. 又

$$M = E = \{(0,x_2),(x_1,0),x_i \in \mathrm{R}, i=1,2\},$$

对应于每个 c, 有 $V^{-1}(c)\cap M=\{(0,\pm c),(\pm c,0)\}$. 因为 (2.2.27) 的每个解是有界的, 所以当 $t\to\infty$ 时, 这些解趋向于这些平衡点 $(0,\pm c)$, $(\pm c,0)$ 中的某一个.

例 2.2.7 考虑系统

$$\ddot{x}+a\dot{x}+2bx+3x^2=0, \tag{2.2.28}$$

其中 $a>0$, $b>0$. 将 (2.2.28) 化为等价的方程组形式

$$\begin{cases} \dot{x}=y, \\ \dot{y}=-2bx-ay-3x^2. \end{cases} \tag{2.2.29}$$

易知平衡点为 $(0,0)$ 及 $(-\frac{2}{3}b,0)$. 置

$$V(x,y)=\frac{1}{2}y^2+bx^2+x^3,$$

即它取为 (2.2.29) 的能量函数, 则有 $\dot{V}|_{(2.2.29)}=-ay^2$. 易知 V 是 R^2 上的 (2.2.29) 的 Lyapunov 函数. 使 $\dot{V}=0$ 的点集为 x 轴, 即

$$E=\{(x,0),x\in\mathrm{R}\},$$

E 中最大不变集 $M=\{(0,0),(-\frac{2}{3}b,0)\}$, 即 M 为平衡点集. 设 $G=\{(x,y)\mid V(x,y)<\frac{4}{27}b^3\}$, 显见 G 是 R^2 中的开集, 由于 $\dot{V}|_{(2.2.29)}\leqslant 0$, 所以 G 是正不变的. 将 G 分成两个集合 G_1 与 G_2. G_1 定义为 G 中包含原点以及点 $(-\frac{2}{3}b,0)$ 的右边部分的点 $\{(x,0):0\geqslant x>-\frac{2}{3}b\}$ 的集合, G_2 定义为 G 中包含点 $(-\frac{2}{3}b,0)$ 的左边的那个无界部分的集合, 注意 G_1 和 G_2 均不包含点 $(-\frac{2}{3}b,0)$, 且 G_1,G_2 的选取使 G_1 与 G_2 是正不变的, 从 G_1 或 G_2 出发的任意的 (2.2.29) 的解都不趋向于平衡点 $(-\frac{2}{3}b,0)$. 定理 2.2.19 的条件中 G 取为 G_1 集合是成立的, 此时对应的 M 就是原点, 这是因为对应定理中 M 的 M_1 为两个平衡点, $\overline{M_1\cap G}$ 即为原点, 且 M^0 是紧致的, V 在 M^0 的边界即原点取值为常数值, 故由定理 2.2.19 得原点是渐近稳定的吸引子且 G_1 是原点的吸引域. 对 G_2 中出发

的解，平衡点 $(-\frac{2}{3}b, 0)$ 是不稳定的，或在有限时间或在 $t \to \infty$ 时解趋向于无穷而不趋向于 $(-\frac{2}{3}b, 0)$.

3. 具有延迟的泛函微分方程情形 Lasalle 不变性原理的应用

考虑 n 维自治的具有延迟的泛函微分方程

$$\dot{x}(t) = f(x_t), \tag{2.2.30}$$

这里 $x : [-r, a) \to \mathrm{R}^n$, $C = C([-r, 0], \mathrm{R}^n)$ 是连续函数 $\phi : [-r, 0] \to \mathrm{R}^n$ 的空间，$\|\phi\|_C = \sup\limits_{t \in [-r, 0]} \|\phi(t)\|$. 在 C 空间中的收敛性是 $[-r, 0]$ 上的一致收敛性. $x_t : [-r, 0] \to \mathrm{R}^n$, $x_t(\theta) = x(t + \theta)$, $-r \leqslant \theta \leqslant 0$, $0 \leqslant t < a$. 设 X 是 C 中的一个开集，$f : X \subset C \to \mathrm{R}^n$ 是连续的.

(2.2.30) 的初值条件为

$$x_0 = \phi \in C. \tag{2.2.31}$$

假定 $[-r, \omega(\phi))$ 是初值问题 (2.2.30), (2.2.31) 的解的最大存在区间，记 $I(\phi) = [0, \omega(\phi))$. 以下总假设 f 满足一定条件，使 (2.2.30) 在 $I(\phi)$ 中的解存在且唯一.

定义 2.2.17 记 $\chi = \{(t, \phi) : t \in I(\phi), \phi \in X\} \subset \mathrm{R}^+ \times X$, **解映射** $\pi : \chi \to X$ 定义为 $\pi(t, \phi) = x_t$, $\pi(t, \phi)(0) = x_t(0) = x(t)$.

定义 2.2.18 函数 $\psi \in C$ 称为 $\pi(t, \phi)$ 的 **正的极限点**, 是指存在 $t_n \in I(\phi)$, 使得当 $n \to \infty$ 时有 $t_n \to \omega(\phi)$ 及 $\pi(t_n, \phi) \to \psi$. **正极限集** $\Omega(\phi)$ 是 $\pi(t, \phi)$ 的正的极限点的集合. 注意 $\pi(t, \phi)$ 可能有正极限点在 χ 的边界上.

定义 2.2.19 集合 $H \subset X$ 称为关于方程 (2.2.30) 是 **不变的**，指 $I(\phi)$ 对每个 $\phi \in H$ 为 $[0, \infty)$, 又对所有 $t \geqslant 0$, $\pi(t, H) = H$.

实际上解映射 $\pi(t, \phi)$ 有如下的性质:

(1) 或者 $I(\phi) = [0, \infty)$ 或者 $\Omega(\phi) \cap X$ 是空集;

(2) $\pi(0, \phi) = \phi$;

(3) 由 $s,t\in \mathrm{R}^+=[0,\infty)$ 及 $s+t\in I(\phi)$ 可得 $s\in I(\pi(t,\phi))$ 及 $\pi(s,\pi(t,\phi))=\pi(s+t,\phi)$.

定义 2.2.20 设 $x(t)$ 是初值问题 (2.2.30), (2.2.31) 的解，定义 $\gamma^+(\phi)=\pi(I(\phi),\phi)=\{x:t\in I(\phi)\}$ 为从 ϕ 出发的 **正轨道**. 如果 $\gamma^+(\phi)$ 是紧致的且 $\gamma^+(\phi)\subset X$, 则称 $\gamma^+(\phi)$ 为 **预紧致** 的，即 $\gamma^+(\phi)$ 为相对于 X 是预紧致的，这等价于 $\gamma^+(\phi)$ 相对于 C 是预紧致的且 $\Omega(\phi)\subset X$, 且有如下的结论：若 $x(t)$ 是预紧致的，则 $I(\phi)=[0,\infty)$.

(2.2.30) 的正极限集 $\Omega(\phi)$ 有如下性质：若 $x(t)$ 是 (2.2.30) 的预紧致的解，则 $\Omega(\phi)$ 在 X 中非空紧致不变且连通，又当 $t\to\infty$ 时有 $x_t\to\Omega(\phi)$.

定义 2.2.21 定义 $V:X\to \mathrm{R}$, V 沿 (2.2.30) 解的全导数为

$$\dot V(\phi)=\liminf_{t\to 0^+}\frac{1}{t}[V(\pi(t,\phi))-V(\phi)].$$

设 G 是 X 的子集，如果 V 在 X (或 $\overline{G}$) 上是连续的，且对所有 $\phi\in G$ 有 $\dot V(\phi)\leqslant 0$, 那么称 V 为 (2.2.30) 在 G 上的 **Lyapunov 泛函**.

若 V 是 Lyapunov 泛函，则定义

$$E=\{\phi:\dot V(\phi)=0,\phi\in\overline{G}\}.$$

类似地定义 M 是 E 中的最大不变集.

定理 2.2.20 (具延迟的泛函微分方程的 Lasalle 不变性原理) 如果 V 是 (2.2.30) 的在 G 上的 Lyapunov 泛函，又 (2.2.30) 的解 $x(t)$ 是预紧致的，且使得对所有 $t\geqslant 0$, x_t 保持留在 G 中. 那么对某些 c, 当 $t\to\infty$ 时成立 $X_t\to M\cap V^{-1}(c)$, 即

$$\lim_{t\to\infty}\inf\{\|x_t(\phi)-\phi\|_C:\phi\in M\cap V^{-1}(c)\}=0.$$

定理 2.2.21 如果 V 是 $U_a=\{\phi\in C:V(\phi)<a\}$ 上的 Lyapunov 泛函，且存在常数 $k=k(\sigma)$, 使得当 $\phi\in U_a$ 时就有

$|\phi(0)| < k$, 则对任意 $\phi \in U_a$, 当 $t \to \infty$ 时都有 $x_t(\phi) \to M$.

定理 2.2.22 若存在泛函 $V : C \to \mathrm{R}$ 为连续的, $V(0) = 0$, 又存在非负连续非减函数 $u(r)$ 和 $w(r)$, 当 $r \to \infty$ 时, $u(r) \to \infty$, $w(0) = 0$, 使得

(1) $u(|\phi(0)|) \leqslant V(\phi)$;

(2) $\dot{V}(\phi) \leqslant -w(|\phi(0)|)$,

则方程 (2.2.30) 的零解稳定, 所有解都为有界. 如果 $w(r)$ 为定正, 则当 $t \to \infty$ 时所有解趋于零.

注 2.2.6 定理 2.2.20 中要求 (2.2.30) 的解 $x(t)$ 是预紧致的, 而实际上只要求 (2.2.30) 的解 $x(t)$ 是有界的, 且在 G 的边界上 x_t 没有正的极限点, 因为可以证明如下的命题:

若 $x(t)$ 是 (2.2.30) 的有界解, 又 x_t 在 G 的边界上没有正的极限点, 则 $x(t)$ 是预紧致的.

下面举例说明不变性原理是如何应用的.

例 2.2.8 考虑具延迟的微分差分方程

$$\dot{x}(t) = ax^3(t) + bx^3(t-r). \tag{2.2.32}$$

注意到当 $r = 0$ 时, 若 $a + b < 0$, 则 (2.2.32) 的零解是整体渐近稳定的; 若 $a + b > 0$, 则 (2.2.32) 除了平凡解外其它所有解都是无界的.

置

$$V(\phi) = -\frac{1}{2a}\phi^4(0) + \int_{-r}^{0} \phi^6(\theta)\mathrm{d}\theta,$$

由

$$V(x_t) = -\frac{1}{2a}x^4(t) + \int_{t-r}^{t} x^6(s)\mathrm{d}s$$

得

$$\dot{V}(\phi)|_{(2.2.32)} = -\left[\phi^6(0) + \frac{2b}{a}\phi^3(0)\phi^3(-r) + \phi^6(-r)\right].$$

以下假定 $|b| \leqslant |a|$, 显见

$$\dot{V}(\phi)|_{(2.2.32)} \leqslant -\left(1-\frac{b^2}{a^2}\right)\phi^6(0) \leqslant 0,$$

即 $V(\phi)$ 为 C 上的 Lyapunov 泛函. 若 $a<0$, 则由 $V(\phi) \geqslant \dfrac{1}{2|a|}\phi^4(0)$ 及定理 2.2.22 知 (2.2.32) 的每个解是有界的, 从而是预紧致的以及零解是稳定的.

(1) $a<0$ 且 $|b|<|a|$ 的情形. 易知 $E=\{\phi\in C:\phi(0)=\phi(-r)=0\}$, $M=\{C$ 空间中的零元素 $\}$. 由定理 2.2.22 知方程 (2.2.32) 的零解是全局渐近稳定的.

(2) $a<0$ 且 $b=a$ 的情形. 易知 $E=\{\phi\in C:\phi(0)=-\phi(-r)\}$, 又因 M 为 E 中的最大不变集, 所以 M 必须是满足 $x(t)=-x(t-r)$ 的解的初值集合. 由 $b=a$ 得 $\dot{x}(t)=0$, 即 $x(t)=c$ (常数). 从而 $\phi(\theta)=c$, $\theta\in[-r,0]$. 但由 $\phi(-r)=-\phi(0)$, 得 $c=-c$, 所以 $c=0$, 故 $M=\{C$ 中的零元素 $\}$, 所以由定理 2.2.22 知方程 (2.2.32) 的零解是全局渐近稳定的.

(3) $a<0$ 且 $b=-a$ 的情形. 易知 $E=\{\phi\in C:\phi(0)=\phi(-r)\}$, $M=\{[-r,0]$上的常数函数$\}$. 又知正极限集 $\Omega(\phi)$ 是连通的, $\Omega(\phi)\subseteq M\cap V^{-1}(c)=\{$有限个$[-r,0]$上常函数$\}$. 这是根据 $V(\phi)$ 在 $\phi=a$ (常数) 时是 a 的 6 次多项式, 对 $V(a)=c$ 有限多个 a 满足. 故 $\Omega(\phi)$ 为单点集, 即 (2.2.32) 的每个解当 $t\to\infty$ 时趋向于一个常数, 此时每个常数均是 (2.2.32) 的平衡点.

(4) $a>0$ 且 $|b|<a$ 或 $b=a$ 的情形. 此时集合 $G=\{\phi\in C:V(\phi)<0\}$ 是非空及正不变的, $M=\{0\}$, 因此 M 在 G 的边界上, 每个从 G 中出发的解 x_t 经过有限时间或 $t\to\infty$ 趋向于无界, 即不能趋向于 $M=\{0\}$, 也就是零解是不稳定的.

2.2.3 奇异吸引子及混沌动力学分析

近年来, 混沌奇异吸引子已成为人们研究的热点之一, 下面将通过一些实例来介绍混沌奇异吸引子的特征、判别及应用. 区分一维离散动力系统、高维离散动力系统、连续动力系统几种情形进行

详细的讨论.

1. 一维离散动力系统的混沌动力学

一维离散动力系统中的抛物线映射是混沌动力学的典型例子，通过对这个实例的分析可以了解耗散系统的混沌奇异吸引子的特征.

抛物线映射是由虫口模型

$$x_{n+1} = \mu x_n(1 - x_n) \tag{2.2.33}$$

引出的，记 $F_\mu(x) = \mu x(1 - x)$.

以 (2.2.33) 中 $\mu = 4$ 为例， $y = 4x(1 - x)$ 是一个完全确定性的迭代方程，但是迭代的结果呈现出如下的 3 个特征的复杂动力学行为:

(1) 迭代的结果对初值的极端敏感依赖性;

(2) 迭代的结果似乎是随机的;

(3) 迭代的结果是长期不可预测的.

初始值取 3 组值，仅在小数点后第 8 位上分别取为 0, 1, 2, 这种误差在一般的计算机中是不可避免的，具体迭代的结果已在表 2.1.1 中列出.

下面考虑抛物线映射 $F_\mu = \mu x(1 - x)$ 的一些普适性质:

(1) 通向混沌的周期倍分叉的 Feigenbaum 道路

先来看一下混沌发生前，即 $1 < \mu < \mu_\infty = 3.5699$ 时稳定吸引子的变化情况. 见图 2.2.2, 把横坐标取为参数 μ, 纵坐标取为 ξ, 对某一横坐标参数 μ, 描绘出对应的稳定的平衡态的值 ξ 的轨迹. 当 μ 从 $\mu_1 = 1$ 逐渐增加到 $\mu_2 = 3$ 时，稳定的平衡态是一个值；而从 μ_2 变到 μ_3 时，稳态为 2 个值；从 μ_3 变到 μ_4 时，稳态为 $4 = 2^2$ 个值；一般地，从 μ_k 变到 μ_{k+1} 时，稳态为 2^{k-1} 个值； μ 变到 μ_∞, 倍周期分叉直到 2^∞ 周期即非周期稳态；当 $\mu_\infty \leqslant \mu \leqslant 4$ 时，几乎在 $[0, 1]$ 中的每个值均有可能变为初始迭代后出现的稳态值，即迭代可能达到的值几乎充满整个 $[0, 1]$. 当然也有空白的带状子区域，这是所谓的窗口，在窗口中又会重新呈现出从 μ_1 到 4 的但

尺寸比例缩小的相似结构.

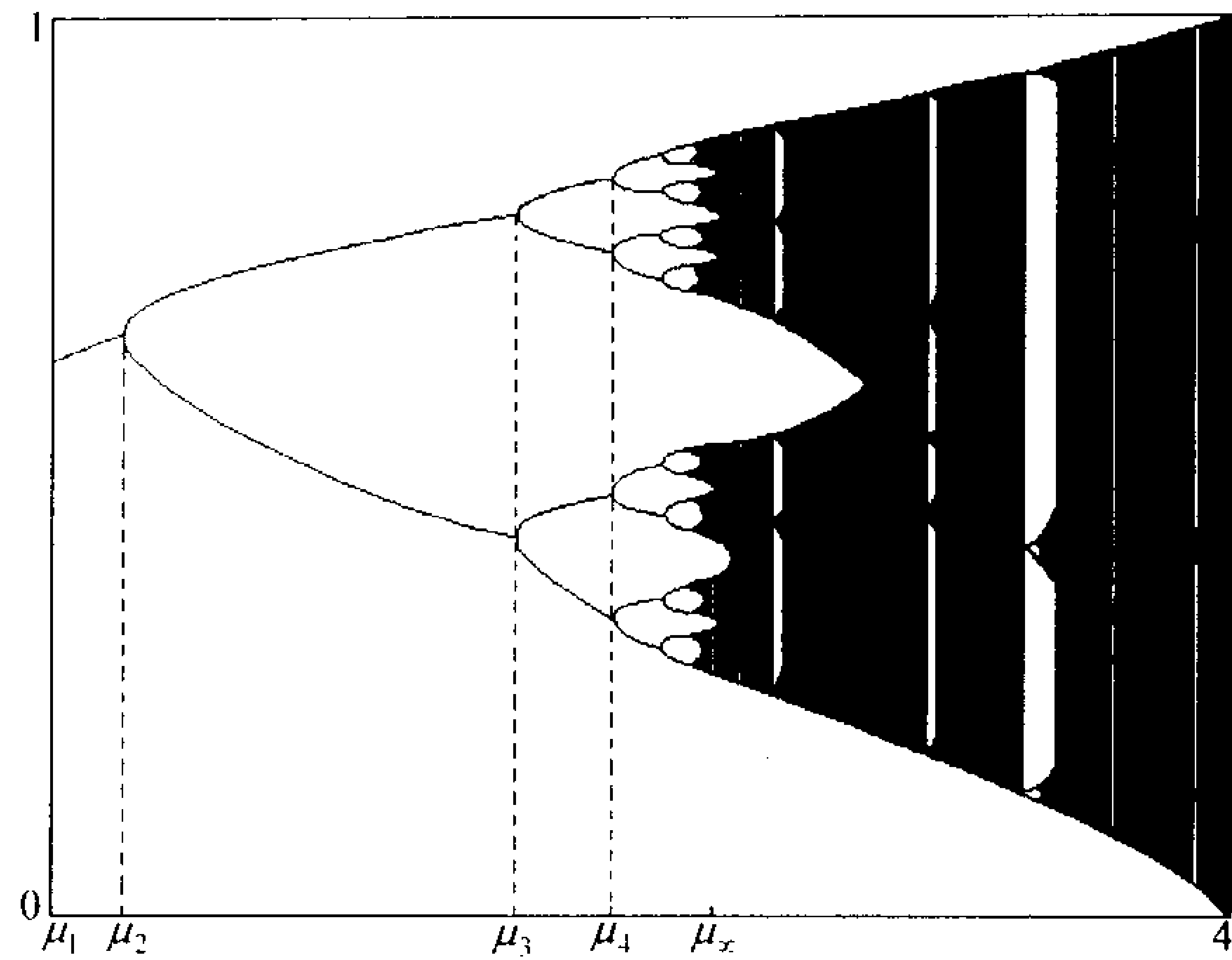


图 2.2.2　抛物线映射的倍周期分叉通向混沌的道路

一般地，对于一维的区间映射

$$x_{n+1} = f(\mu, x_n), \quad n = 1, 2, \cdots, x_n \in I, \tag{2.2.34}$$

当 f 满足一定的条件，例如单峰函数或仅有单一最大值的函数，Feigenbaum 发现了一条通向混沌的典型道路，一个系统一旦发生了周期倍分叉，必导致混沌. 称这条道路为 **Feigenbaum 道路**.

(2) Feigenbaum 道路的两个普适常数 δ, α

抛物线映射的倍周期的分叉点序列 μ_n 是一个无穷序列，它有一个极限 $\mu_\infty = 3.569945672\cdots$, 且这个序列的下列形式的差商有极限 δ, 它是一个普适常数

$$\delta = \lim_{m\to\infty} \frac{\mu_m - \mu_{m-1}}{\mu_{m+1} - \mu_m} = 4.6692016091029909\cdots,$$

称为 **Feigenbaum 常数**. 具体的分叉值及差商值见表 2.2.1.

表 2.2.1　抛物线映射周期倍分叉的分叉点及差商值

n	分叉情况	μ_m	$\dfrac{\mu_m-\mu_{m-1}}{\mu_{m+1}-\mu_m}$
1	1 分为 2	3	
2	2 分为 4	3.449489743	4.751446
3	4 分为 8	3.544090359	4.656251
4	8 分为 16	3.564407266	4.668242
⋯	⋯	⋯	⋯
9	256 分为 512	3.569943176	4.669
10	512 分为 1024	3.569945137	4.669
⋯	⋯	⋯	⋯
∞	2^∞ 周期解 → 混沌解	3.569945672	4.6692016

在 ξ 轴上相邻周期解的周期之间的距离按因子

$$\alpha = 2.502907875095892 84\cdots$$

缩减，这是第二个普适常数. 这两个常数是一切周期倍分叉的道路中所共有的，但不是说不同的映射，如果它们都是周期倍分叉道路通向混沌的话，那么都取同样的 δ 和 α 的值，只是说都存在具有这种性质的常数 δ 和 α, 但映射不同的话，算出来的 δ 和 α 的值也是不一样的.

下面介绍区间映射 (2.2.34) 为混沌的数学涵义.

(1) 在 R. Devany 意义下的数学涵义

设 $f: I \to I$, I 为实轴上一区间，那么具有下面 3 个要素的映射 f 是混沌的:

(i) 长期不可预测性.　f 对初始条件敏感依赖，如存在 $\delta > 0$, 对任意 $x \in J$ 和 x 的任何邻域 N, 存在 $y \in N$ 和 $n \geqslant 0$, 使得 $|f^n(x) - f^n(y)| > \delta$.

(ii) 不可分解性. f 是拓扑传递的, 如对任意一对开集 $U, V \subset I$, 存在 $k>0$, 使得 $f^k(U) \cap V \neq \emptyset$.

(iii) 有一定规律性. f 的周期点在 I 中稠密.

上述 3 条混沌的数学涵义是 1986 年由 R. Devany 出版的书“混沌动力系统引论”中给出的 (参见本章参考文献 [1]).

(2) 在 Li–Yorke 意义下的数学涵义

区间映射 f 如果满足如下 3 个条件, 那么 f 是混沌的:

(i) f 具有任意正整数周期的周期点, 即对任意自然数 n, 有 $x \in I$, 使得 $f^n(x) = x$;

(ii) 存在不可数集 $K \subset I$, 使得对任意 $x, y \in K$ 有

$$\overline{\lim_{n\to\infty}} |f^n(x) - f^n(y)| > 0,$$
$$\underline{\lim_{n\to\infty}} |f^n(x) - f^n(y)| = 0;$$

(iii) 对每一 $x \in K$ (不可数集) $\subset I$ 及周期点 y 有

$$\overline{\lim_{n\to\infty}} |f^n(x) - f^n(y)| > 0.$$

进一步地有一个十分有趣的结果: 若 $f: I \to I$ 是连续的, 且有周期 3 的点, 则 f 是上述 Li–Yorke 涵义下的混沌映射, 也就是周期 3 意味着混沌.

区间映射 f 为混沌的, 上述两种涵义有一定的联系及差异.

以下介绍一维离散动力系统为混沌的判别方法.

方法 1　直接从上述混沌两种涵义 (Devany 意义与 Li–Yorke 意义) 来判别.

例 2.2.9　考虑 $f_{S^1}: S^1 \to S^1$, S^1 为单位圆周, $f_{S^1}(\theta) = 2\theta$. 下面验证 f_{S^1} 满足 Devany 意义下的混沌涵义.

在 f_{S^1} 的迭代下任意两点 θ_1 与 θ_2 之间的角距离被加倍, 取 $\delta = 10^{-1}$, 必有 $k>0$, 使得

$$|f_{S^1}^k(\theta_1) - f_{S^1}^k(\theta_2)| = 2^k|\theta_1 - \theta_2| > 10^{-1}.$$

即证得 f_{S^1} 满足对初值敏感依赖性.

因为 S^1 上的任何一段小弧都可以由 f_{S^1} 的某一 k 次迭代最终扩展以覆盖整个 S^1, 即 S^1 上任何一段弧都能被覆盖, 所以 f_{S^1} 的拓扑传递性质满足.

最后, f_{S^1} 的周期点即满足 $f_{S^1}^n(\theta) = \theta + 2k\pi$, $0 \leqslant k \leqslant 2^n$, k 为正整数, $\theta = \dfrac{2k\pi}{2^n - 1}$, 即 $(2^n - 1)$ 次单位根, 易见 f_{S^1} 的周期点在 S^1 上稠密.

由于上述验证的混沌三要素满足, 故 f_{S^1} 为混沌的.

方法 2 因为与标准的混沌映射 f (例如单位圆上的倍角映射 f_{S^1}) 拓扑共轭的映射 g 也是混沌的, 所以只要找到 g 与 f 之间的同胚映射 h 就能判别 g 映射是混沌的.

首先给出拓扑共轭的定义.

定义 2.2.22 设 $f: A \to A$ 及 $g: B \to B$ 为两个映射, 如果存在一同胚 $h: A \to B$, 使得 $h \circ f = g \circ h$, 则称 f 和 g 是 **拓扑共轭** 的, 同胚 h 被称为 **拓扑共轭**.

其次介绍如下的定理:

定理 2.2.23 对于拓扑共轭的各种映射, 它们的动力性态是完全等价的. 这些动力性态包括不动点与周期点的个数、类型以及混沌映射的 Devany 意义下的三要素等.

例 2.2.10 证明 F_4 满足 Devany 意义的混沌三要素条件.

为此建立例 2.2.9 中的映射 f_{S^1} 与 F_4 之间的同胚关系, 即找到使得 f_{S^1} 和 F_4 拓扑共轭的同胚 h.

参见图 2.2.3, 注意在定理 2.2.23 中, 拓扑共轭的条件可以适当放宽一些, 例如下面寻找的 h 仅使 f_{S^1} 与 F_4 是半拓扑共轭的, 此时定理 2.2.23 的结论同样成立.

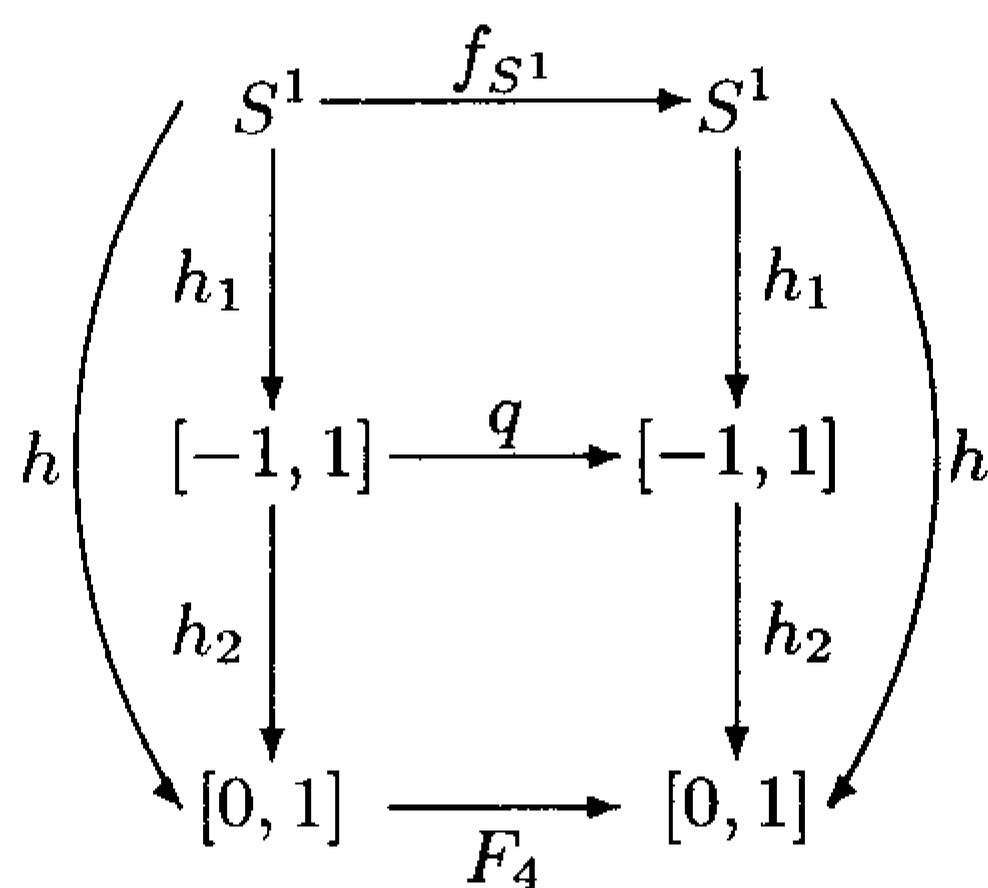


图 2.2.3 f_{S^1} 与 F_4 两映射半拓扑共轭示意图

定义 $h_1(\theta) = \cos\theta$, $h_1 : S^1 \to [-1,1]$, $h_2(t) = \frac{1}{2}(1-t)$, $h_2 : [-1,1] \to [0,1]$. 再定义 $h = h_2 \circ h_1$, 即 $h : S^1 \to [0,1]$, $h(\theta) = \frac{1}{2}(1-\cos\theta)$. 此时 h_1 不是一个同胚，它在大多数点上是二对一的. 现在两个映射 $f_{S^1} : S^1 \to S^1$, $f_{S^1}(\theta) = 2\theta$, $F_4 : [0,1] \to [0,1]$, $F_4(y) = 4y(1-y)$. 又定义 $q(x) = 2x^2 - 1$, 显然 $q : [-1,1] \to [-1,1]$.

可以证明， q 通过 h_1 共轭于 f_{S^1}, 即

$$h_1 \circ f_{S^1} = \cos 2\theta = 2\cos^2\theta - 1 = q \circ h_1(\theta),$$

又易证明， $F_4 \circ h_2 = h_2 \circ q$, 即

$$F_4 \circ h_2(x) = 4\frac{1}{2}(1-x)(1-\frac{1}{2}(1-x)) = 1 - x^2,$$

$$h_2 \circ q(x) = \frac{1}{2}(1-(2x^2-1)) = 1 - x^2.$$

所以，可证得

$$F_4 \circ h_2 \circ h_1 = h_2 \circ q \circ h_1 = h_2 \circ h_1 \circ f_{S^1},$$

将 $h = h_2 \circ h_1$ 代入公式得

$$F_4 \circ h = h \circ f_{S^1},$$

从而由 f_{S^1} 满足 Devany 混沌三要素可以推得 F_4 也满足 Devany 的混沌三要素，即知 F_4 在 Devany 意义下是混沌的映射.

方法 3 用某些特征量的数值计算 (例如 Lyapunov 指数、相关维数、熵、复杂度等) 来判别映射是否具有混沌.

(1) Lyapunov 指数

混沌的产生定性地可以理解为拉伸压缩的整体的平均效果最终是拉伸，但整个系统又是有界的，用一个特征量 Lyapunov 指数来描述. 这是一种常用的判别混沌的数字特征. 下面给出它的数学定义以及对吸引子类型的判据.

假设间隔长度为 ε 的两个点 x_0 与 $x_0+\varepsilon$, 经过一次迭代映射 f 的作用变到新的两个点 $f(x_0)$ 与 $f(x_0+\varepsilon)$, 它们被拉伸的平均因子是 $\mathrm{e}^{\lambda(x_0)}$, 即像点间距离为 $\varepsilon\mathrm{e}^{\lambda(x_0)}$. 若 N 次迭代，则被拉伸的平均因子为 $\mathrm{e}^{N\lambda(x_0)}$ (见图 2.2.4).

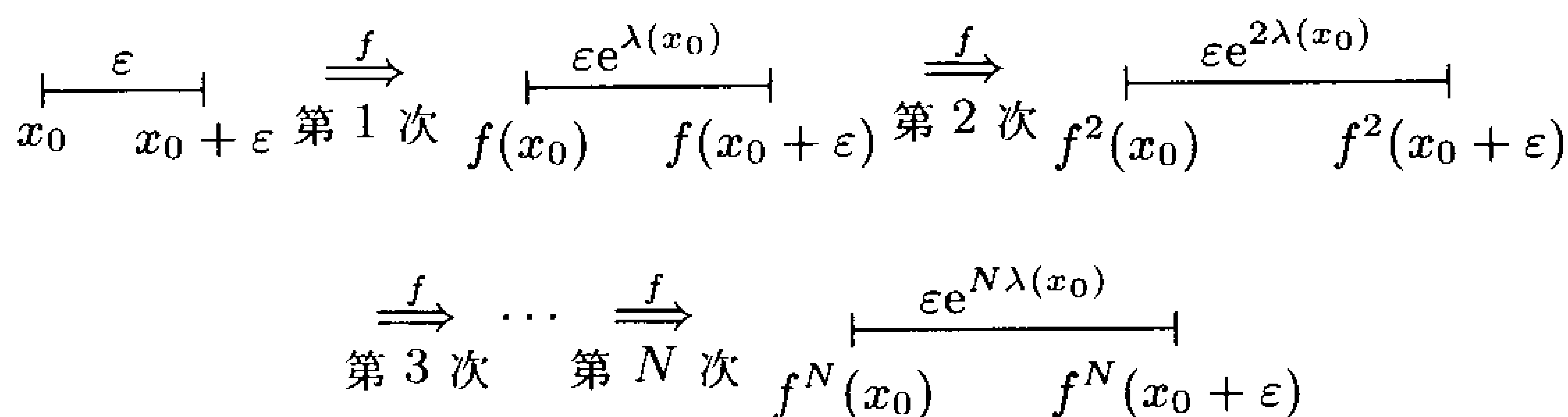


图 2.2.4　Lyapunov 指数的定义示意图

用数学形式表示为

$$|f^N(x_0+\varepsilon)-f^N(x_0)|=\varepsilon\mathrm{e}^{N\lambda(x_0)}.$$

当 $\varepsilon\to 0$ 和 $N\to+\infty$ 时，它就导致了 $\lambda(x_0)$ 的正确表达形式. 现定义 Lyapunov 指数

$$\begin{aligned} l(x_0) &= \lim_{N\to\infty}\lim_{\varepsilon\to 0}\frac{1}{N}\ln\left|\frac{f^N(x_0+\varepsilon)-f^N(x_0)}{\varepsilon}\right| \\ &= \lim_{N\to\infty}\frac{1}{N}\ln\left|\frac{\mathrm{d}f^N(x_0)}{\mathrm{d}x}\right|. \end{aligned} \tag{2.2.35}$$

注意到求复合函数导数的链式法则，有

$$\left.\frac{\mathrm{d}f^2(x)}{\mathrm{d}x}\right|_{x=x_0}=\left.\frac{\mathrm{d}}{\mathrm{d}x}f[f(x)]\right|_{x=x_0}=f'[f(x_0)]f'(x_0)=f'(x_1)f'(x_0),$$

其中 $x_1 \triangleq f(x_0)$.

一般地有

$$\left.\frac{\mathrm{d}f^N(x)}{\mathrm{d}x}\right|_{x=x_0} = f'(x_{N-1})f'(x_{N-2})\cdots f'(x_0),$$

其中 $x_i \triangleq f(x_{i-1})$ $(i = 1, 2, \cdots, N-1)$, 所以 Lyapunov 指数可以定义为

$$\begin{aligned} l(x_0) &= \lim_{N\to\infty} \frac{1}{N} \ln \left| \prod_{i=0}^{N-1} f'(x_i) \right| \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} \ln |f'(x_i)|. \end{aligned} \tag{2.2.36}$$

由 (2.2.35) 及 (2.2.36) 可见 $l(x_0)$ 与 x_0 有关, 但是对于如 F_μ 那样的一维的单峰映射, 可以证明 $l(x_0)$ 对于 $x_0 \in I$ 时为恒定的, 记为 l. 一维离散动力系统只有 1 个 Lyapunov 指数, 对于二维以上的离散动力系统, Lyapunov 指数的个数与维数相同, 那时数字特征指标就是取最大的 Lyapunov 指数 l. 由 l 大于零、等于零或小于零来判别映射的混沌特征, 正的 Lyapunov 指数表示 f 映射迭代运动轨道在每个局部点 x 处, 在长时间多次迭代 N 次的平均意义下为拉伸的, 即不稳定的, 相邻轨道被指数 e^{lN} 分离, 由于轨道在整体的稳定因素 (像有界耗散等) 作用下反复折叠, 形成混沌奇异吸引子; 负的 Lyapunov 指数表示轨道在局部是稳定的, 形成了周期吸引子.

例 2.2.11　求 F_4 的 Lyapunov 指数 l.

现在用下面的式子求 Lyapunov 指数

$$L = \lim_{n\to\infty} |f'(x_0)f'(x_1)\cdots f'(x_n)|^{\frac{1}{n+1}}. \tag{2.2.37}$$

此时的混沌判据要修改为 $L > 1$, 周期吸引子的判据为 $L < 1$. 在抛物线映射的周期窗口的参数分叉是始点处的 $L = 1$.

在 F_4 的情形取初值 $x_0=0.49$, 此时

$$L=4\lim_{n\to\infty}|(1-2x_0)(1-2x_1)\cdots(1-2x_n)|^{\frac{1}{n+1}},$$

其中 $x_j=F_4(x_{j-1})=4x_{j-1}(1-x_{j-1})\ (j=1,2,\cdots,n)$.

实际上计算下面序列

$$y_n=4|(1-2x_0)(1-2x_1)\cdots(1-2x_n)|^{\frac{1}{n+1}},$$

对 $n=0,1,2,\cdots,16$ 计算得相应的 y_n 值为

$$0.08, 0.57, 1.08, 1.50, 1.80, 1.99, 1.84, 1.98, 1.93,$$
$$2.00, 1.65, 1.78, 1.89, 1.97, 1.99, 1.95, 2.00$$

得 $L=2$. 可见 F_4 是混沌的. 理论上 L 是序列 y_n 的极限. 实际操作是观察 y_n, 计算其相邻几个值相差比较小时就可以停止算下去. 取其中的一个值即可以, 当然这是不很严格的, 但对实际应用还是可取的.

一般地对 F_μ $(3\leqslant\mu\leqslant 4)$ 有 Lyapunov 指数的图 (见图 2.2.5), 此时以 l 即取对数的形式来计算 Lyapunov 指数.

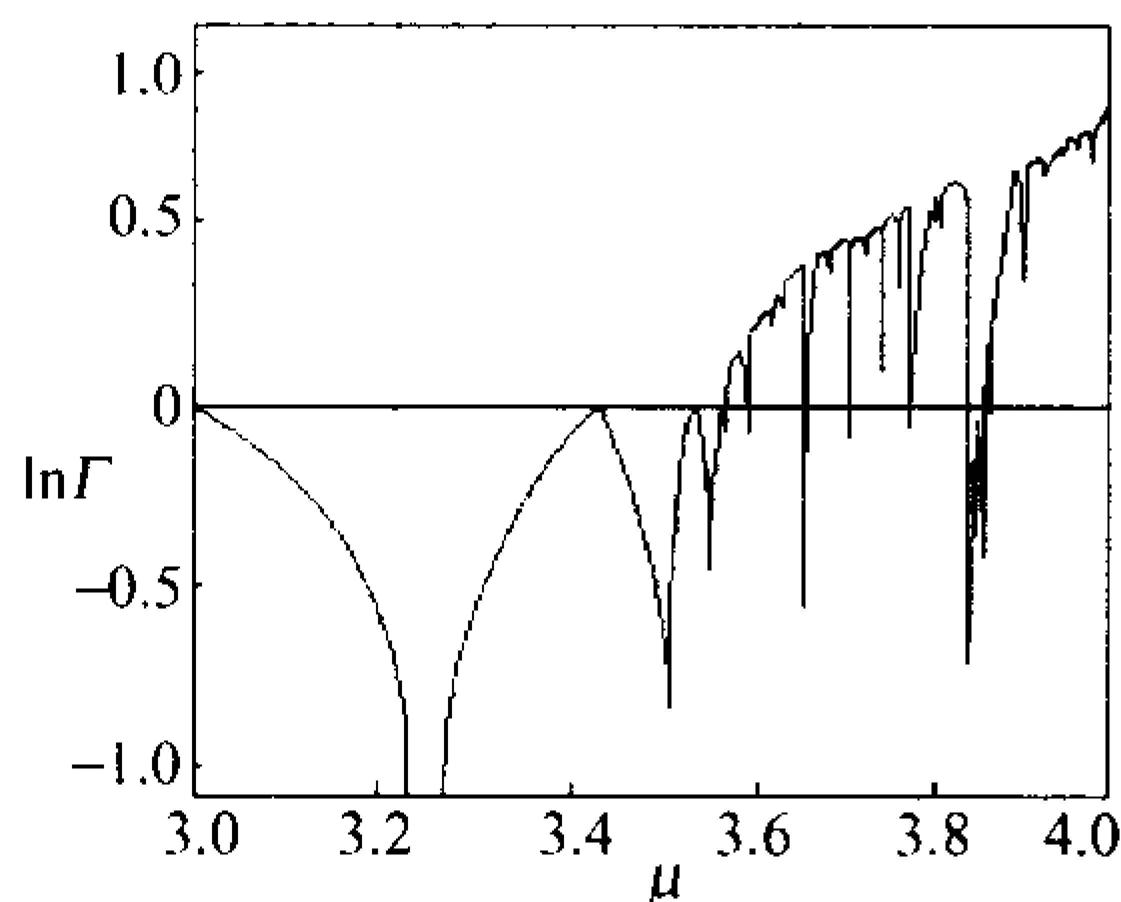


图 2.2.5　以参数 μ 为横坐标, Lyapunov 指数 L 的对数值为纵坐标

从上面 F_μ 的 Lyapunov 指数图可见, 当 μ 从 3 变到 $\mu_\infty=3.5699\cdots$ 时, 对应倍周期的稳定的轨迹, Lyapunov 指数均为负

数，对应 2^n 周期的分支点，例如 $\mu \approx 3.45$ 时过分支点，它相应于 2^n 稳定周期轨道与 2^{n+1} 稳定周期轨迹的交叉点边缘，此时 Lyapunov 指数为零. 也有一些超稳定点，例如 $\mu = 3.50$ 时，它相应于 Lyapunov 指数为负无穷大，此处轨迹的收敛速度比指数速度还快. 当 $\mu \in (\mu_\infty, 4]$ 时， F_μ 进入混沌状态， Lyapunov 指数总体上为逐渐增大的正数，但又不时地为无限多个具有周期行为的"窗口"所中断. 总之，该图的 Lyapunov 指数的变化很好地反映了 Logistic 映射的动力学特性.

有时 (2.2.36) 中的对数用以 2 为底的对数 $\log_2$ 形式，此时记 Lyapunov 指数为 λ. 这种表示有如下好处，它意味着每次迭代所损失的信息量，而这个量可以用对数来表示. 也就是说，对于任何 μ 值及任何精确的初始条件， λ 值反映了系统失去预测能力的平均速度. 例如 $\mu = 4$ 时， $\lambda = 1.0$, 它表示系统按每次迭代失去 1 位信息量的速度逐渐降低预测能力. 如果初始条件包含 10 位信息，那么经过 8 次迭代后只剩下 2 位信息量， 10 次迭代后系统就完全失去预测能力. 最后我们指出从映射的各种等效形式出发得到相同的 Lyapunov 指数.

例 2.2.12 求帐篷映射 T_r $(r > 0)$ 的 Lyapunov 指数，这里

$$T_r(x) = r\left(1 - 2\left|\frac{1}{2} - x\right|\right). \tag{2.2.38}$$

$r = 1$ 的情形见图 2.2.6.

下面讨论 T_1 的 Lyapunov 指数. 由图 2.2.6(b) 可见，$\left|\dfrac{\mathrm{d}}{\mathrm{d}x}T_1^n(x)\right| = 2^n$, "几乎所有的" 点 x_0 和 $x_0 + \varepsilon$ 之间的距离在 n 次迭代后随 n 作指数增长，并且其 Lyapunov 指数为 $L = 2$ 或 $l = \ln 2$, 注意它与 x_0 无关. 一般地，对 T_r 可以求得其 Lyapunov 指数为 $L = 2r$ 或 $l = \ln(2r)$. 当 $r \in [0, \frac{1}{2})$ 时， $L < 1$, 故 T_r 为周期的，实际上一切点都被吸引到 $x^* = 0$ 点. 当 $r > \frac{1}{2}$ 时， $L > 1$, 故 T_r 为混沌的.

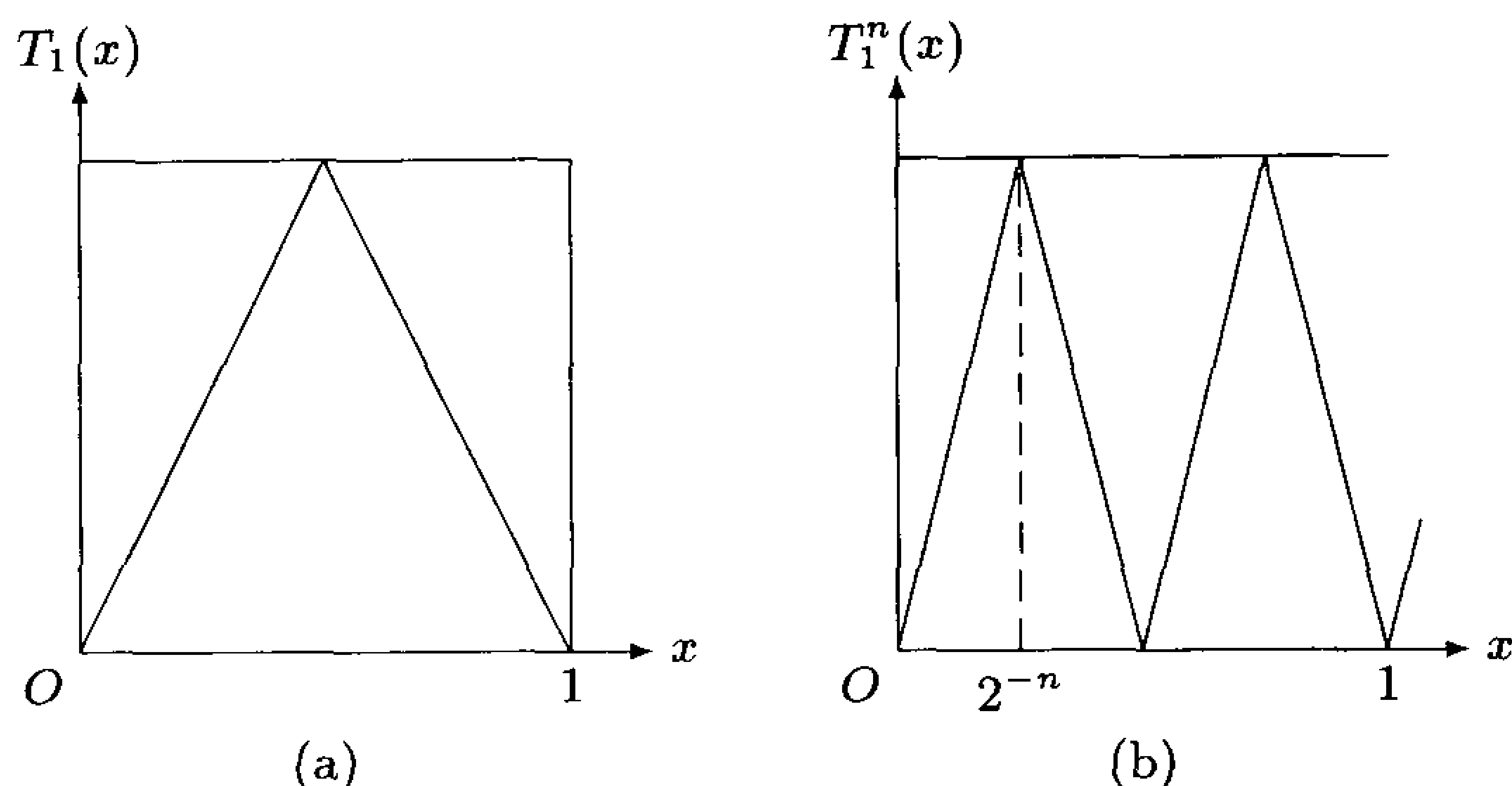


图 2.2.6　(a) 帐篷映射 T_1; (b) 第 n 次迭代 T_1^n

(2) 测度熵

前面给出的 Lyapunov 指数大于零时，一维离散映射的轨道呈现局部不稳定性，使相邻轨道以指数速率分离. 现在引进另一个数值指标——**测度熵**, 又称为 **Kolmogorov–Sinai 熵**. 下面从信息角度来分析这个数值指标. 虽然测度熵的数学定义以及具体的计算比较复杂，但是它与 Lyapunov 指数有一定的数量上的对应关系，从而也就可以从信息的角度来理解 Lyapunov 指数为正的含义. 这些内容对于混沌动力学在信息科学的应用是富有启发性的.

从信息产生或丧失的角度去理解 Lyapunov 指数. 假如两个初始点很靠近以致在一段时间里不能靠测量来区分两条轨道，而在过了相当一段时间后，相邻轨道被分离而能够加以区分，也就是说随着时间的推移，混沌运动产生信息，这个信息量就是可以区分的不同轨道的数目 N. 此时 N 随时间指数增长

$$N \propto \mathrm{e}^{kt},$$

其中常数 k 刻画信息产生的速率，实际上就是测度熵. 从另一方面也可以这样来理解，如果我们掌握关于初值的精确信息，对一维动力学情形，即初始点在 $[0,1]$ 的精确位置的信息，随着运动过程这些信息逐渐丢失，轨道变得越来越远离最初的点或者轨道变得不能同来自其它初值的某些轨道区分. 上面常数 k 的倒数，具有时

间的量纲，给出从精确初值出发可以进行预测的时间. 测度熵 K 的定义如下:

$$K = -\lim_{\Delta t \to 0} \lim_{\varepsilon \to 0} \lim_{m \to \infty} \sum_{i_1 \cdots i_m} p(i_1, i_2, \cdots, i_m) \times \ln p(i_1, i_2, \cdots, i_m), \tag{2.2.39}$$

这里 ε 是球或箱的尺寸，设轨道在时刻 t 时处于第 i_1 号箱中，在时刻 $t+\Delta t$ 时处于第 i_2 号箱中, $\cdots\cdots$, 在时刻 $t+(m-1)\Delta t$ 时处于第 i_m 号箱中的概率为 $p(i_1, i_2, \cdots, i_m)$, 它原则上可以从数值实验或实际观测中求得.

测度熵 K 的数值是判别运动性质的重要指标，对于规则运动， $K=0$; 对于随机运动， $K=\infty$; 对于混沌运动，对应有限的正 K 值.

测度熵 K 与 Lyapunov 正指数 l 的关系, 在一维情形， $K=l$, 这意味着正的 Lyapunov 指数也具有上述讲述的从两个不同角度理解的信息产生与丢失的含义.

上述的例子 —— 抛物线映射具有一维相空间和一维参数空间, 它展示了一条倍周期分叉的通向混沌的道路. 下面将介绍的圆周映射, 具有一维相空间和二维参数空间, 它通向混沌的道路比抛物线映射的 Feigenbaum 道路要复杂，它在一个复杂的模式中表现为周期的、拟周期的和混沌的运动, 它提供了从拟周期性通向混沌的道路.

例 2.2.13 考虑一维的圆映射

$$\theta_{n+1} = f(\theta_n) \equiv \theta_n + \Omega + \frac{K}{2\pi}\cos(2\pi\theta_n) \pmod 1, \tag{2.2.40}$$

上式角度 θ_n 的运动描述了从拟周期性到混沌的转变， θ_n 的模数是 1, 即 $\cos[2\pi(k+\theta_n)] = \cos 2\pi\theta_n$ $(0 \leqslant \theta_n < 1)$. 映射 $f: S \to S$, S 为圆周， K 与 $\varOmega$ 是 f 映射的两个参数，因此有的书或论文中改记 f 为 $f_{K\varOmega}$. 下面讨论随两个参数 $K, \varOmega$ 的变化得到的各种动力学行为.

(1) 不动点

取 N 为整数，不动点的方程为

$$N = \Omega + \frac{K}{2\pi}\cos 2\pi\theta. \tag{2.2.41}$$

显见该方程可能得到 1 个、 2 个或多个稳定的或不稳定的不动点的解. 特别地，对 $N = 0$, 有

$$\theta = \frac{1}{2\pi}\arccos\left(-\frac{2\pi\Omega}{K}\right). \tag{2.2.42}$$

将 (2.2.40) 右边求导，可求出稳定性条件

$$-1 < 1 - K\sin(2\pi\theta) < 1. \tag{2.2.43}$$

(2) 临界点，即在相空间中映射 f 的一阶导数为零的点

对圆映射 (2.2.40), 当 $K < 1$ 时，对所有的 θ, 有

$$f'(\theta) = 1 - K\sin(2\pi\theta) > 0,$$

即没有临界点; 当 $K = 1$ 时，称为参数空间中的临界线; 当 $K > 1$ 时，从几何上易看出，存在两个临界点，即

$$1 - K\sin(2\pi\theta) = 0.$$

在 $[0, 2\pi]$ 中有两个根，当映射有两个临界点时，称为 **双程式** 的. 多于 1 个临界点的存在具有更为重要的结果，这就是在参数空间的给定点. 有可能存在 1 个以上的吸引子，此时有这样的结果，同时存在的稳定周期轨迹的数目不大于 2.

(3) Arnol'd 舌头

现在来分析 K–Ω 两参数区域中的一些楔形区域上的既有趣又复杂的动力学性质. 先置 $K = 0$, $\Omega = \dfrac{p}{k}$, p 和 k 为质数，此时 (2.2.40) 变为

$$\theta_{n+1} = \theta_n + \frac{p}{k} \pmod 1. \tag{2.2.44}$$

从 θ_0 开始，迭代 k 次，则有

$$\theta_k = \theta_0 + p(\text{mod } 1) = \theta_0. \tag{2.2.45}$$

显见在参数空间中经过一点的所有轨道都是周期 k 轨道，而与相空间的起始点无关. 对于给定的 k, 可能有许多选择 p 的方法，只要 Ω 是以 k 为分母的有理数，就一定能找到周期 k 轨迹. 比率 $\frac{p}{k}$ 称为轨迹的卷曲数 (winding number). 更一般地，可以定义对于任意 K 和 Ω, 任何无限长周期轨迹的卷曲数为

$$W = \lim_{n\to\infty} \frac{\theta_n - \theta_0}{n}. \tag{2.2.46}$$

在这个公式中 θ_n 的计算无需考虑模 1 的问题.

从 (2.2.44) 可知，当 $k > 0$ 时，在一个有限宽度的区域中不动点是稳定的. 固定 $k \in (0,1)$, 由数值模拟能得到起始于 $k = 0$ 的每一个有理数 Ω 的一个个楔形区域，这些楔形区域通常被称为 Arnol'd 舌头 (见图 2.2.7).

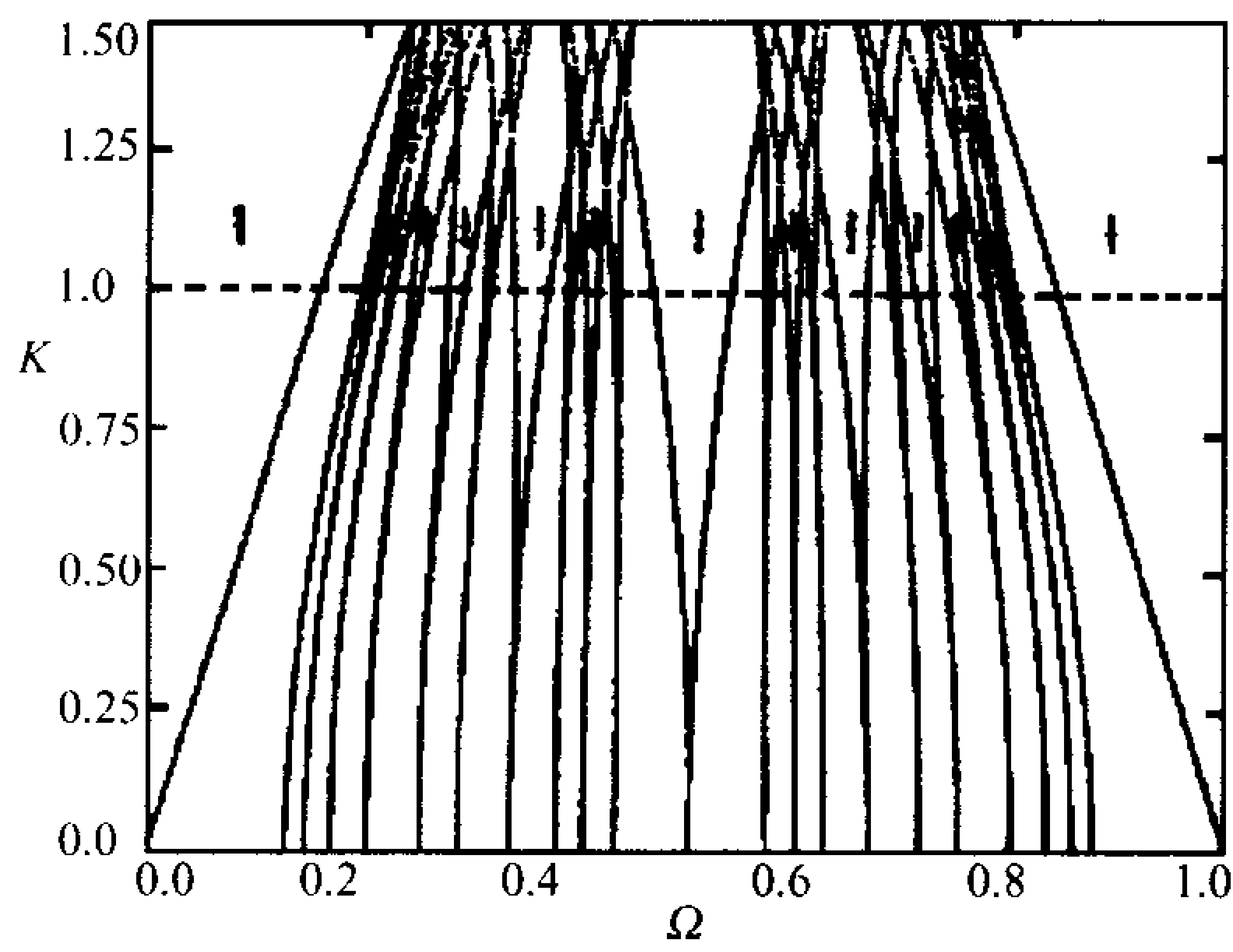


图 2.2.7　Arnol'd 舌头

通常，把有理卷曲数看作两个竞争频率是很有用的，其中一个是系统频率，另一个是外部驱动的频率. 在参数空间的一个有限区域里，卷曲数锁定于一个不变的有理数的现象称作 **锁相** 或 **锁频**. 在许多动力学系统里，这是一种很重要的运动.

圆映射的 Lyapunov 指数，当 $k = 0.9, 1.0, 2.7$ 时依赖于 Ω 变化，这从图 2.2.8, 图 2.2.9, 图 2.2.10 可见.

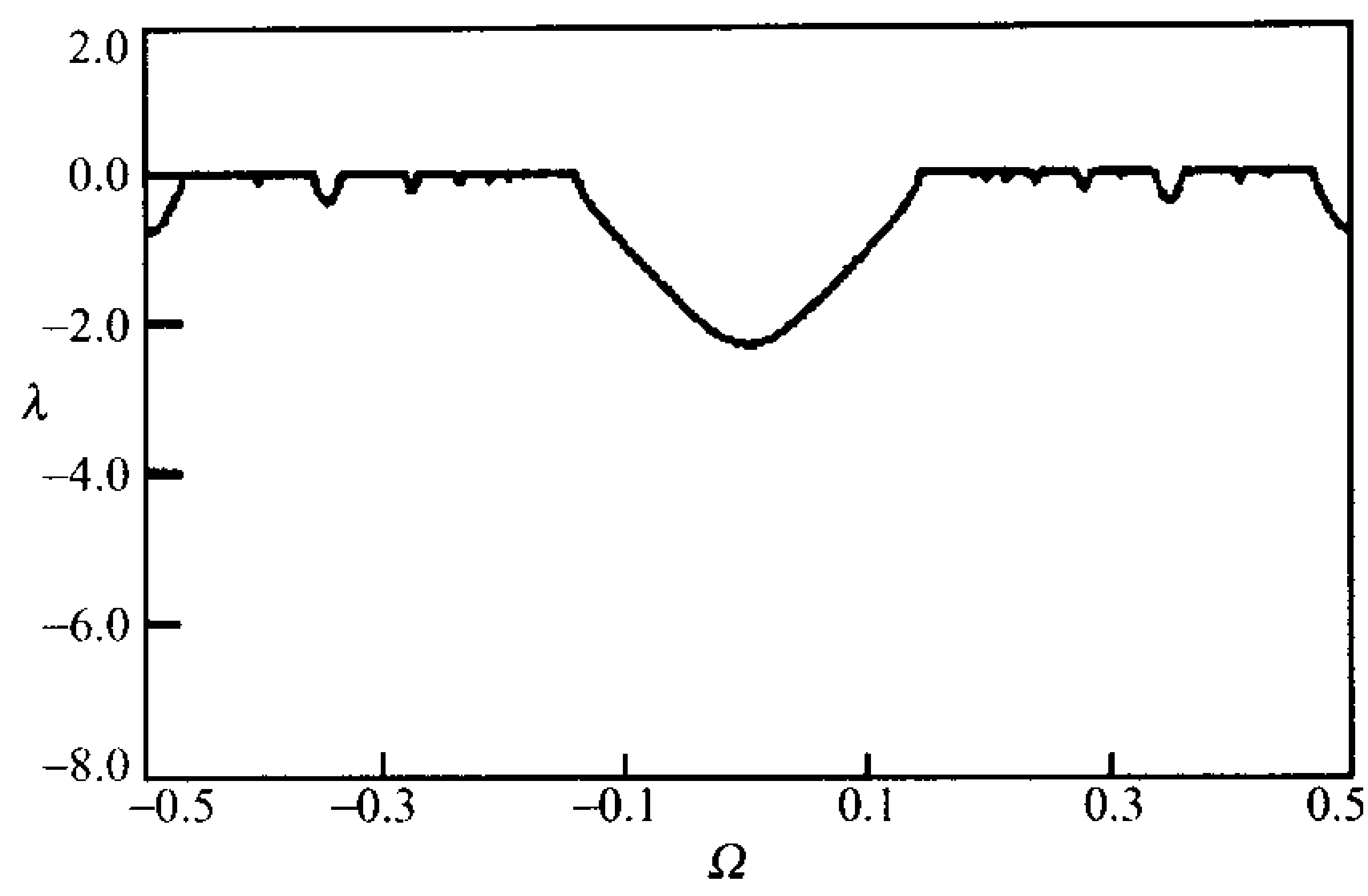


图 2.2.8　　$K = 0.9$ 时，圆映射的 Lyapunov 指数

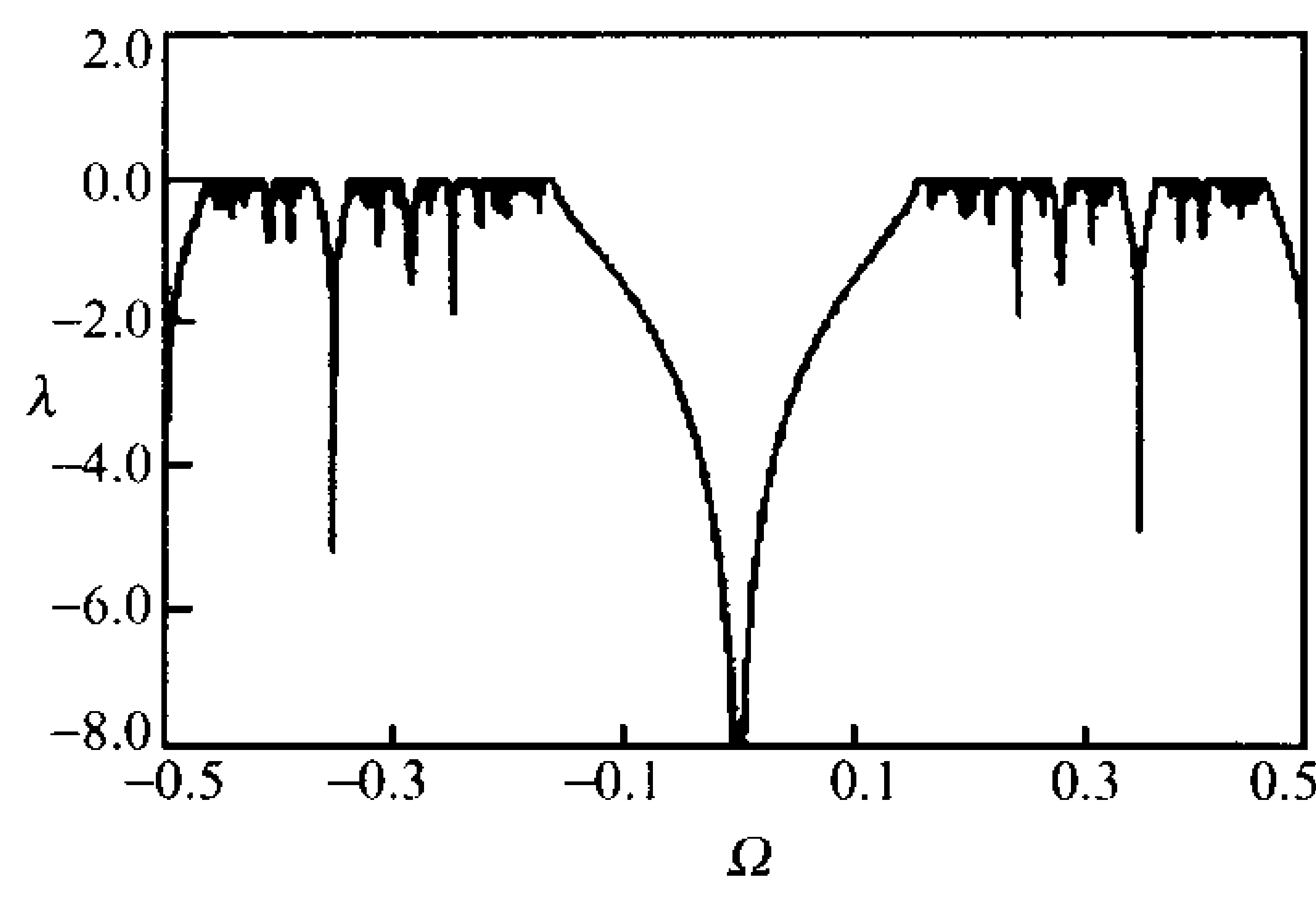


图 2.2.9　　$K = 1.0$ 时，圆映射的 Lyapunov 指数

Lyapunov 指数为负的情形对应于稳定的周期轨道，Lyapunov 指数为零的情形对应于拟周期轨道，而 Lyapunov 指数为正的情形对应于混沌轨道.

在 $K < 1$ 时，没有混沌轨迹，而得到了所谓的 Arnol'd 舌头，在那里运动是模锁的，即卷曲数 W 是无理数，即在 K–Ω 平面上的两个区域，模锁和非模锁都是有限的. 这是一个向拟周期轨迹过渡的过程.

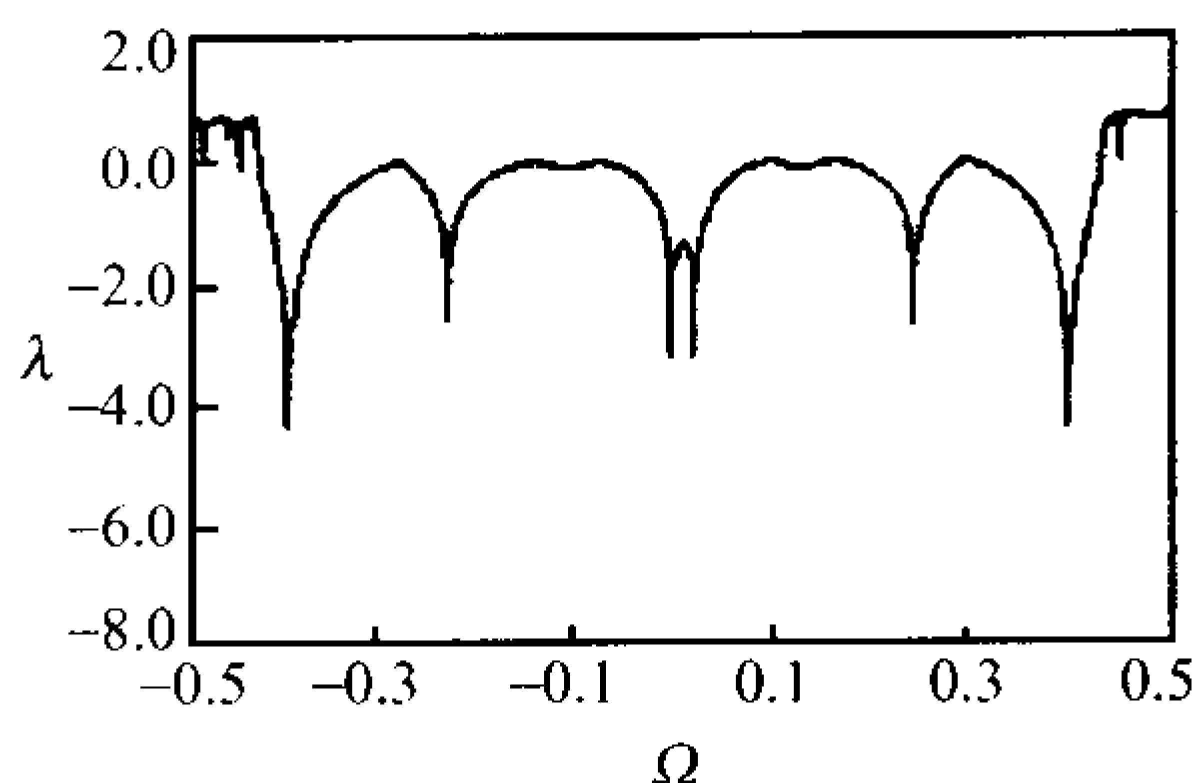


图 2.2.10　$K = 2.7$ 时，圆映射的 Lyapunov 指数

在 $K = 1$ 时， Arnol'd 舌头以这样一种方式一起移动，剩余的非模锁的 Ω 间隔区域 "空洞" 构成一个 Cantor 集合.

在 $K > 1$ 时，映射 $f_{K\Omega}$ 变得不可逆，可能产生混沌行为，但混沌和非混沌区域在 K–Ω 平面上稠密地交织在一起，即存在两种可能性，要么过渡到混沌，要么过渡到另一周期轨迹，其周期性决定于 Ω.

对 $K < 1$, $K = 1$, $K > 1$ 三种情形的圆映射的相图的示意图见图 2.2.11.

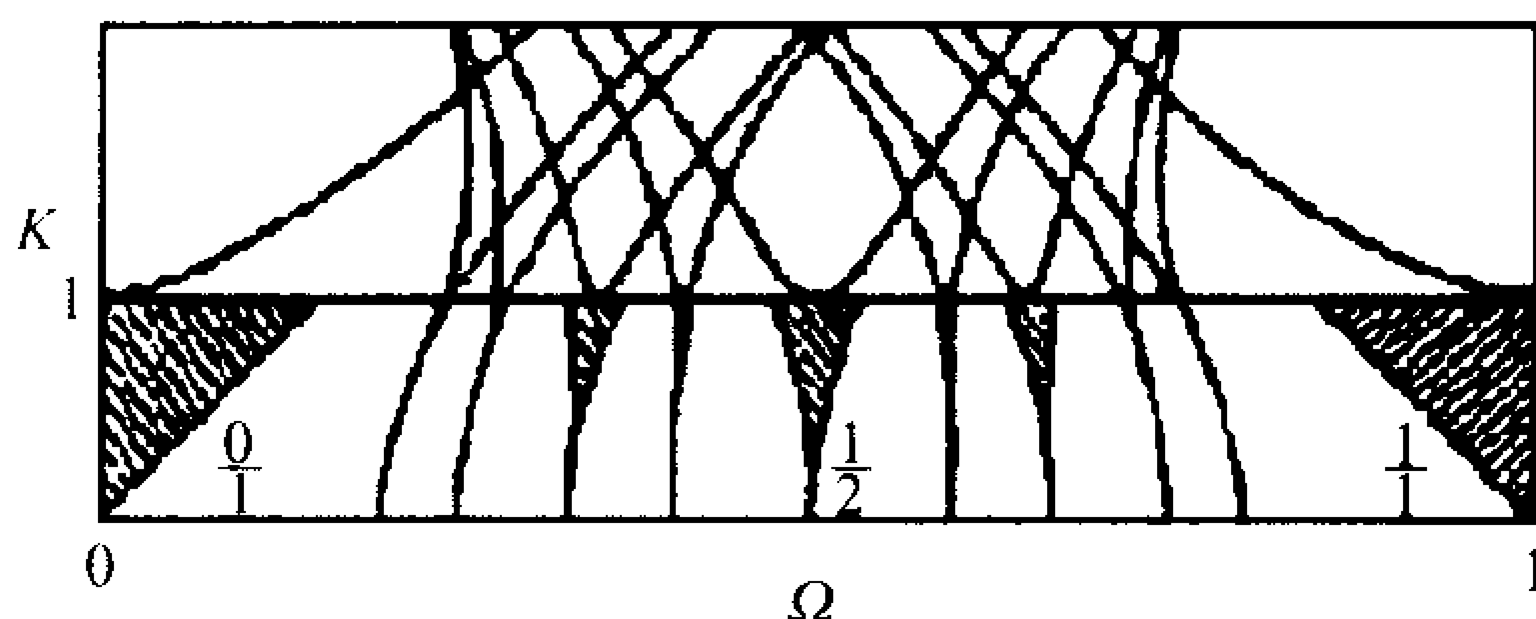


图 2.2.11　圆映射的相图 (示意图)

与从倍周期分叉过渡到混沌道路的抛物线映射类似，从拟周期性向混沌过渡的道路的圆映射的普适特征为

(1) $f(\theta)$ 具有 $f(\theta + 1) = 1 + f(\theta)$ 的性质;

(2) 对 $|K| < 1$, $f(\theta)$ 和它的倒数 $f^{-1}(\theta)$ 存在且可微，即 $f(\theta)$ 是一个微分同胚;

(3) 对 $K = 1$, $f^{-1}(\theta)$ 成为不可微，但对 $|K| > 1$, $f(\theta)$ 不存在

唯一的倒数.

2. 高维离散动力系统的混沌动力学

高维系统的奇异吸引子呈现出很复杂的动力学行为，目前的研究进展与一维系统比较差距很大，二维离散动力系统还有一些理论上的结果，高维系统仅仅有一些实际上可操作的数量特征的数值计算方法，现简单介绍如下.

以二维映射为例，考虑映射 $f:(x,y)\mapsto(X(x,y),Y(x,y))$, 记其 Jacobi 矩阵为

$$J(x,y)=\begin{pmatrix}\dfrac{\partial X}{\partial x} & \dfrac{\partial X}{\partial y}\\[2ex] \dfrac{\partial Y}{\partial x} & \dfrac{\partial Y}{\partial y}\end{pmatrix}.$$

由初始点 $P_0(x_0,y_0)$ 出发，逐次得映像的轨道为 $P_1(x_1,y_1),P_2(x_2,y_2),\cdots,P_n(x_n,y_n),\cdots$, 各次映射对应的 Jacobi 矩阵为 $J_n\triangleq J(x_n,y_n)$ $(n=1,2,\cdots)$. 设矩阵 $[J_0J_1\cdots J_n]$ 的特征值的模为 j_1 和 j_2 (不妨设 $j_1\geqslant j_2$). 对 j_1 与 j_2 开 $n+1$ 次根，这两个值的极限

$$L_i=\lim_{n\to\infty}(j_i)^{\frac{1}{n+1}}\quad(i=1,2)\tag{2.2.47}$$

是映射 f 的两个 Lyapunov 指数. 对于耗散系统，由于面积在映射下是收缩的，故可知 $L_1L_2<1$. 判别 f 是否为混沌的判据是 L_1 与 L_2 中最大的 Lyapunov 指数大于 1, 现在已经规定 $L_1\geqslant L_2$, 故只需判别 L_1 是否大于 1 即可. 当然这里必须说明的是，一般情形下， L_1, L_2 与初始点 P_0 的选取是无关的，具体计算 L_i 时取极限只是理论上的，实际上是给定一个正 ε 小量，对相邻几个 n 求 $(j_i)^{\frac{1}{n+1}}$ 的值，若它们相差小于 ε, 就认为该值为 L_i 的近似值. 有时 Lyapunov 指数的计算式子取为

$$l_i=\lim_{n\to\infty}\frac{1}{n+1}\ln j_i\quad(i=1,2),$$

其中 j_i $(i=1,2)$ 为矩阵 $[J_0J_1\cdots J_n]$ 的两个特征值的模，规定 $j_1\geqslant j_2$, 此时 f 为混沌的判据为 $l_1>0$.

对于二维映射的讨论完全可以类似地推广到 n 维的映射. 在 n 维离散系统中, 存在 n 个 Lyapunov 指数 λ_i $(i=1,2,\cdots,n)$. 如果其中最大的指数 $\lambda_{\max}>0$, 则该系统一定存在混沌. 记矩阵 $[J_0J_1\cdots J_n]$ 的 n 个特征值中模最大的值为 j_1, 则 $\lambda_{\max}=\lim_{n\to\infty}\frac{1}{n+1}\ln j_1$.

对于测度熵 K, 高维情形有如下关系:

$$K\leqslant\sum_{\lambda_i>0}\lambda_i, \tag{2.2.48}$$

其中 λ_i 为正的 Lyapunov 指数. 在实践中等式往往成立, 即在高维系统中, K 量度相空间里一个基本单元 (称之为 **元胞**) 的平均变形, 而且等于正的 Lyapunov 指数的和.

3. 连续动力系统的混沌动力学

连续动力系统的混沌奇异吸引子的典型例子就是 Lorenz 方程 (见例 2.1.8) 的参数取某些值时得到的 Lorenz 奇异吸引子.

Lorenz 方程所描述的系统为耗散系统, 此类系统在运动时, 其相空间容积收缩到维数低于原来相空间维数的吸引子上, 即其运动特征是相空间容积收缩. 随着系统参数的变化, 耗散系统会出现混沌, 这时的吸引子就变为奇异吸引子.

对于一般的 n 维耗散系统

$$\frac{\mathrm{d}x}{\mathrm{d}t}=f(x),\quad x\in\mathrm{R}^n,\quad x_0=x(0), \tag{2.2.49}$$

该系统相空间容积的变化率为

$$\alpha=\sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i}=\mathrm{div}f. \tag{2.2.50}$$

当 $\alpha=0$ 时, 即系统为自治的 Hamilton 系统, 它的特征是能量守恒, 相空间容积保持不变. 对于耗散系统 $\alpha<0$, Lorenz 方程 (2.1.31) 可计算 (2.2.50) 式得

$$\alpha=\frac{\partial}{\partial x}(-\sigma(x-y))+\frac{\partial}{\partial y}(-xz+rx-y)+\frac{\partial}{\partial z}(xy-bz)$$

$$= -(\sigma + 1 + b)$$

是负值，说明 Lorenz 方程代表了一个耗散系统.

由于耗散系统的相空间容积是收缩的，所以 n 维耗散系统的稳态运动将位于一个小于 n 维的超曲面 M 上. 粗略地说， M 就是吸引子. 所谓 M 是 **奇异吸引子** 指它对初值是敏感的，相邻轨道指数分离而局部轨道呈现不稳定性. 又奇异吸引子 M 局限于有限的区域里，就大范围而言，表现为稳定的吸引子，它有自己的吸引域. 若以吸引域内任意一点为初值，则可得到几乎完全相同的奇异吸引子，这表现了 M 的稳定性. M 的空间结构十分复杂，具有无穷层次的自相似结构，这是由轨道的无穷伸展压缩折叠造成的，当然它出现的自治系统的维数最低是 3.

研究连续动力系统的混沌奇异吸引子的方法.

方法 1 通过 Poincaré 截面形成离散的动力系统 Poincaré 映射，对这个离散映射的吸引子分析可以反映连续动力系统的动力学行为.

例 2.2.14 考虑系统

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = x - y - x(x^2 + y^2), \\[2ex] \dfrac{\mathrm{d}y}{\mathrm{d}t} = x + y - y(x^2 + y^2), \end{cases} \tag{2.2.51}$$

将之化为极坐标下的形式

$$\begin{cases} \dfrac{\mathrm{d}r}{\mathrm{d}t} = r(1 - r^2), \\[2ex] \dfrac{\mathrm{d}\theta}{\mathrm{d}t} = 1. \end{cases} \tag{2.2.52}$$

取二维 xOy 平面中的一个一维子空间作为 Poincaré 截面

$$\Sigma = \{(r, \theta) \mid r > 0, \theta = 0\},$$

从 (r_0, θ_0) 出发的 (2.2.52) 的解

$$\begin{aligned} r(t; r_0, \theta_0) &= \left[1 + \left(\frac{1}{r_0^2} - 1\right) \mathrm{e}^{-2t}\right]^{-\frac{1}{2}}, \\ \theta(t; r_0, \theta_0) &= t + \theta_0. \end{aligned}$$

设从 $r_0 > 0$, $\theta = 0$ 的极轴，即从 Σ 上一个初始点 $q_0(r_0, 0)$ 出发，轨线经过时间 $t = 2\pi$ 达到 Σ 上的点 q

$$\begin{aligned} r(2\pi; r_0, 0) &= \left[1 + \left(\frac{1}{r_0^2} - 1\right) \mathrm{e}^{-4\pi}\right]^{-\frac{1}{2}}, \\ \theta(2\pi; r_0, 0) &= 2\pi, \end{aligned}$$

实际上得到 $q\left(\left[1 + \left(\frac{1}{r_0^2} - 1\right) \mathrm{e}^{-4\pi}\right]^{-\frac{1}{2}}, 0\right)$.

显见有如下的映射 P, 把点 q_0 与点 q 建立对应关系，即

$$P(r_0) = \left[1 + \left(\frac{1}{r_0^2} - 1\right) \mathrm{e}^{-4\pi}\right]^{-\frac{1}{2}}. \tag{2.2.53}$$

称该映射 P 为从 Poincaré 截面 Σ 到 Σ 的 **Poincaré 映射**. 如果 $r_0 = 1$, 那么 $P(r_0) = r_0$. 可见 $r = 1$ 为 P 映射的不动点. 易见映射 P 是 1 个一维离散映射，即一维的离散动力系统，而原系统 (2.2.51) 是二维的连续动力系统. 如果 Poincaré 映射 P 是混沌的，那么原连续系统是混沌的，即从 P 的动力学行为可以推得原连续动力系统的相对应的动力学行为.

方法 2 对连续动力系统有相应的混沌动力学的数字特征，例如 Lyapunov 指数是常用来判别混沌的特征. 以下介绍连续动力系统的最大 Lyapunov 指数的计算方法.

选取两个初始条件 x_0 及 y_0, 而其距离 $d = |y_0 - x_0|$, 设 d 值很小. 在一个小的时间间隔 τ 里在微分方程两边关于时间 t 积分，得到的积分的值分别为 x_1 及 y_1, x_1 与 x_0 之间的对应关系，y_1 与 y_0 之间的对应关系均为映射 T^τ, 即

$$x_1 = T^\tau(x_0), \quad y_1 = T^\tau(y_0),$$

这样得到新的两点 x_1, y_1 的距离为 $d_1 = |y_1 - x_1|$. 在 x_1 与 y_1 的连线上选取一个新的点 y_1', 使得 $|y_1' - x_1| = d$. 然后将 x_1 及 y_1' 作为初始点, 也在时间 τ 的小间隔中积分微分方程的两边, 而得到

$$
\begin{aligned}
x_2 &= T^{\tau}(x_1) = T^{2\tau}(x_0), \\
y_2 &= T^{\tau}(y_1').
\end{aligned}
$$

记 $d_2 = |y_2 - x_2|$, 然后在 x_2 与 y_2 的连线上选取 y_2', 使得 $|y_2' - x_2| = d$. 再以 x_2 及 y_2' 作为新的初始点进行上述步骤. 最后得到最大 Lyapunov 指数

$$
l = \ln a = \lim_{m\to\infty}\left(\frac{1}{m\tau}\right)\sum_{i=1}^{m}\ln\left(\frac{d_i}{d}\right), \tag{2.2.54}
$$

其中 m 为上述积分进行的步骤次数. 一般 m 取很大, 例如 10^6, 而 d 取很小, 积分间隔 τ 适当即可, 限制不大. 上述计算方法的步骤可形象地见图 2.2.12.

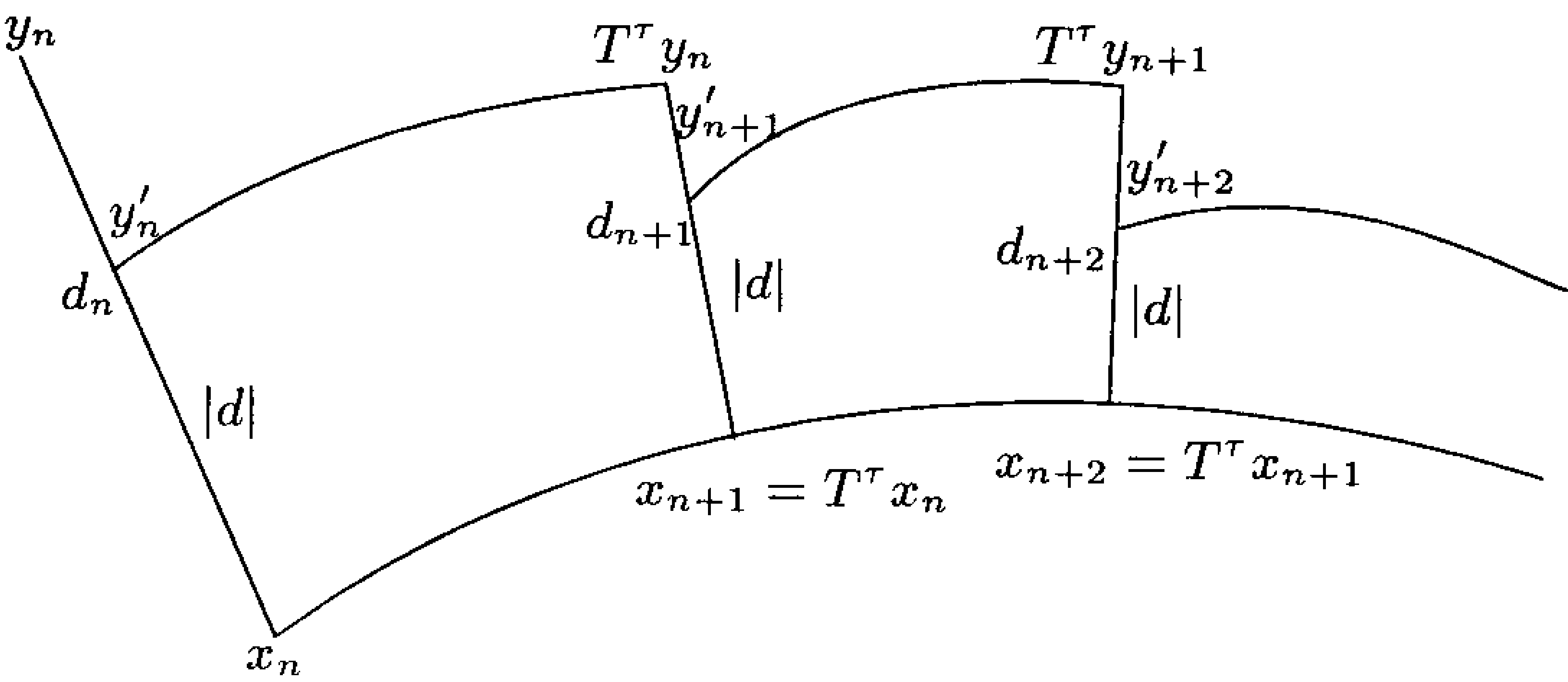


图 2.2.12　计算微分方程组的最大 Lyapunov 指数的简明图解

方法 3　对于二维的自治微分方程加上不定常的强迫项的动力系统, 通过 Poincaré 映射的分析形成的用 Melnikov 函数判别混沌的定量分析方法.

考虑如下的扰动系统

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}+f(x)=\varepsilon g\left(x,\frac{\mathrm{d}x}{\mathrm{d}t},\omega t\right), \tag{2.2.55}$$

其中 g 是 ωt 的周期为 2π 的函数, ε 是扰动项的适当小数.

对于下面的与 (2.2.55) 对应的齐次情形的保守系统

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}+f(x)=0, \tag{2.2.56}$$

假定它有周期解轨道 $x=x^*(t)$, $y=y^*(t)$, 数学上称之为 **同宿轨道解** 或 **异宿轨道解** (以系统的一个鞍点或两个鞍点在该周期闭轨线上来区分).

作如下的 Melnikov 函数

$$M(\tau)=\int_{-\infty}^{+\infty} y^*(t)g(x^*(t),y^*(t),\omega(t-\tau))\mathrm{d}t. \tag{2.2.57}$$

如果函数 $M(\tau)$ 是变号的, 那么系统 (2.2.55) 是混沌的. 实际上从严格的数学论证来看上述结论应该改为有可能是混沌的, (2.2.55) 为混沌的必要条件为 $M(\tau)$ 是变号的. 但是实际上应用时此判据也是很有效的, 故将 $M(\tau)$ 变号作为 (2.2.55) 混沌的判据.

例 2.2.15 考虑如下的系统

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}+\delta\frac{\mathrm{d}x}{\mathrm{d}t}+x-x^3=F\cos\omega t. \tag{2.2.58}$$

引进小量 ε, 将上述系统化为

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}+x-x^3=\varepsilon\left(-\delta'\frac{\mathrm{d}x}{\mathrm{d}t}+F'\cos\omega t\right), \tag{2.2.59}$$

其中 $\delta=\varepsilon\delta'$, $F=\varepsilon F'$, 即 δ 与 F 均是 $O(\varepsilon)$ 的量级. 当 $\varepsilon=0$ 时, 对应齐次的系统是保守系统

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}+x-x^3=0, \tag{2.2.60}$$

它有异宿轨道是

$$x = x^*(t) = \pm\tanh\frac{t}{\sqrt{2}},\quad y = y^*(t) = \pm\frac{1}{\sqrt{2}}\mathrm{sech}^2\frac{t}{\sqrt{2}}.$$

将之代入 Melnikov 函数 (2.2.57) 中得

$$\begin{aligned} M(\tau) &= \int_{-\infty}^{+\infty}\left[F'\cos\omega(t-\tau)\mp\delta'\frac{1}{\sqrt{2}}\mathrm{sech}^2\frac{t}{\sqrt{2}}\right]\frac{\pm 1}{\sqrt{2}}\mathrm{sech}^2\frac{t}{\sqrt{2}}\mathrm{d}t \\ &= -\frac{2\sqrt{2}}{3}\delta' \pm \sqrt{2}\pi\omega F'\frac{\cos\omega\tau}{\sinh\dfrac{\pi\omega}{\sqrt{2}}}. \end{aligned}$$

若 ω 已给定, 只要 δ' 足够大, 就能使上述的 $M(\tau)$ 不变号, 从而防止出现混沌解, 门槛值是

$$R^0(\omega) = \frac{2\sinh\dfrac{\pi\omega}{\sqrt{2}}}{3\pi\omega}.$$

如果强迫力与阻力之比 (设 ε 足够小)

$$\frac{F}{\delta} = \frac{F'}{\delta'} > R^0(\omega),$$

那么系统可能出现混沌解; 如果

$$\frac{F}{\delta} < R^0(\omega),$$

则系统不出现混沌解.

对于系统 (2.2.55) 将相平面 $(x,y) = \left(x, \dfrac{\mathrm{d}x}{\mathrm{d}t}\right)$ 视为 Poincaré 截面 Σ, 也可以构成 Poincaré 映射 P. 由 P 映射的一些动力学性质得原系统 (2.2.55) 的相对应的动力学性质, 它与常微分方程组的解的对应关系见表 2.2.2.

下面介绍通向混沌的阵发道路. 阵发性概念原来是湍流理论中用来描述流场中在层流背景下湍流随机爆发的现象, 表现为层流、湍流相交而使相应的空间域随机地交替. 在混沌的理论中主要是借用阵发性的概念来表示时间域中系统不规则行为和规则行为的随机交替现象.

表 2.2.2　Poincaré 映射的解与常微分方程组的解的对应关系

Poincaré 映射的解	常微分方程组的解
不动点	周期解 (谐波解)
周期点 (周期为整数 m)	周期解，$1/m$ 次谐波解
鞍点的不稳定流形与 (自己的或另一个鞍点的) 稳定流形相交，出现 Cantor 集合型的不变集，且可分析出奇怪不变集合	出现混沌解

1979 年 Pomeau 和 Manneville 对混沌的阵发道路进行了开创性的工作，他们用数值法求解 Lorenz 方程，将 x 分量随时间变化的波形画在图 2.2.13 中.

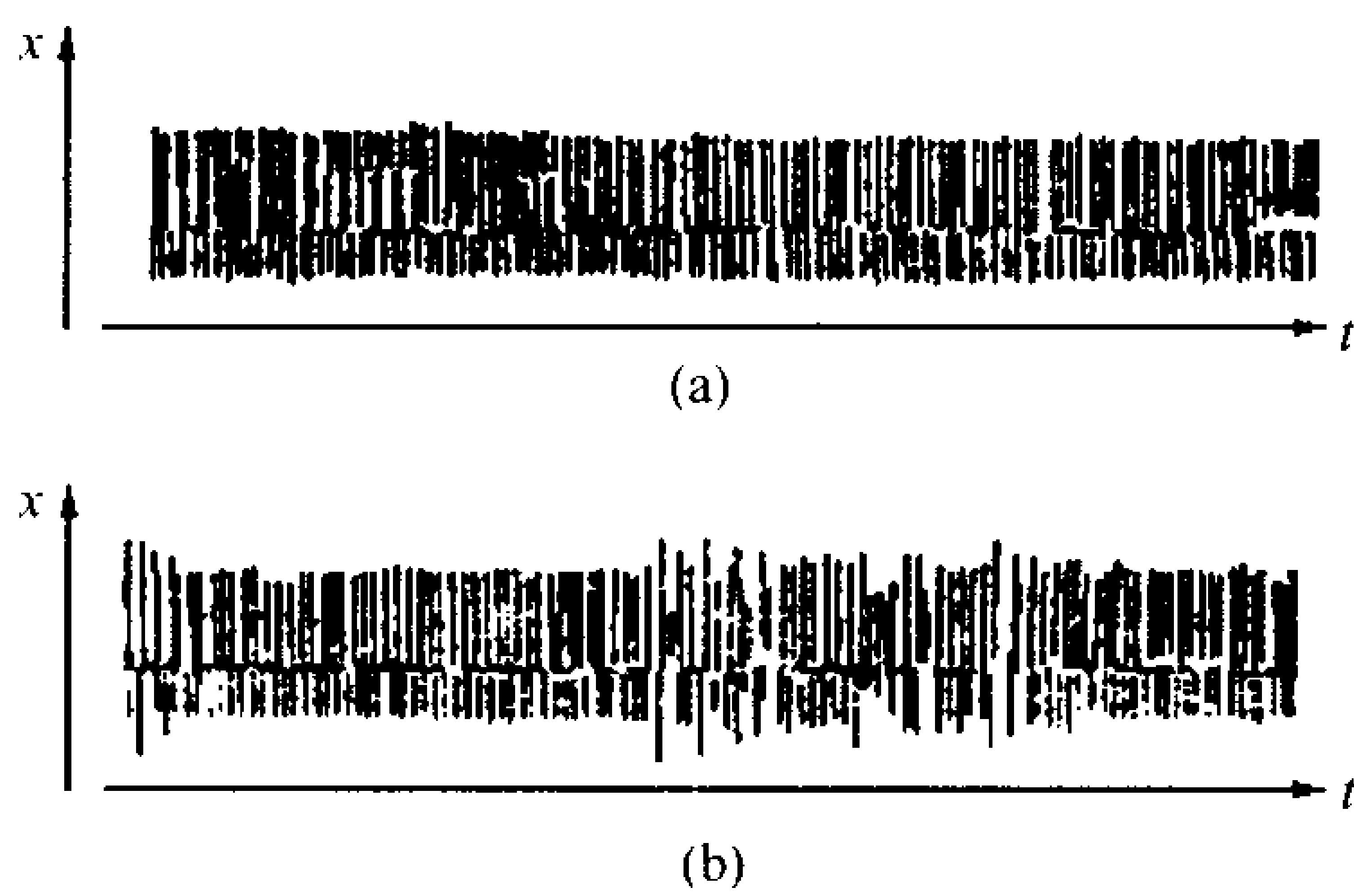


图 2.2.13　Lorenz 方程 x 分量的波形

图 2.2.13(a) 表示当 $r = 166 < r_c = 166.66$ 时 x 轨线呈现出周期运动，对应于 Poincaré 映射中的稳定不动点. 图 2.2.13(b) 表示 $r > r_c$ 时，振荡被不规则的爆发所隔断，对应于 Poincaré 映射中的不动点变成不稳定的. 因为不动点失去其稳定性本质上可以有 3 种方式， Pomeau 和 Manneville 区分了 3 种阵发类型，如表 2.2.3 所示.

当 r 稍大于 $r_c = 166$ 时 Lorenz 模型的 Poincaré 映射见图 2.2.14, 这里画出了 $y(t)$ 穿越 $x = 0$ 平面的值 y_n 的轨迹. 将此图

与表 2.2.3 中的类型 I 比较，易知此时 Lorenz 模型呈现 I 型阵发混沌.

表 2.2.3　阵发的三种类型

类型	特征行为和映射	典型的映射 ($\varepsilon < 0 \to \varepsilon > 0$)	本征值
I	在 $+1$ 穿越单位圆的一个实本征值 $x_{n+1} = \varepsilon + x_n + ux_n^2$		
II	同时穿越单位圆的一对复共轭本征值 $r_{n+1} = (1+\varepsilon)r_n + ur_n^2$ $\theta_{n+1} = \theta_n + \Omega$		
III	在 -1 穿越单位圆的一个实本征值 $x_{n+1} = -(1+\varepsilon)x_n - ux_n^3$		

I 型阵发混沌的转变方式的特征是倒切分岔，其中两个不动点 (一个稳定，一个不稳定) 消失了 (见图 2.2.15).

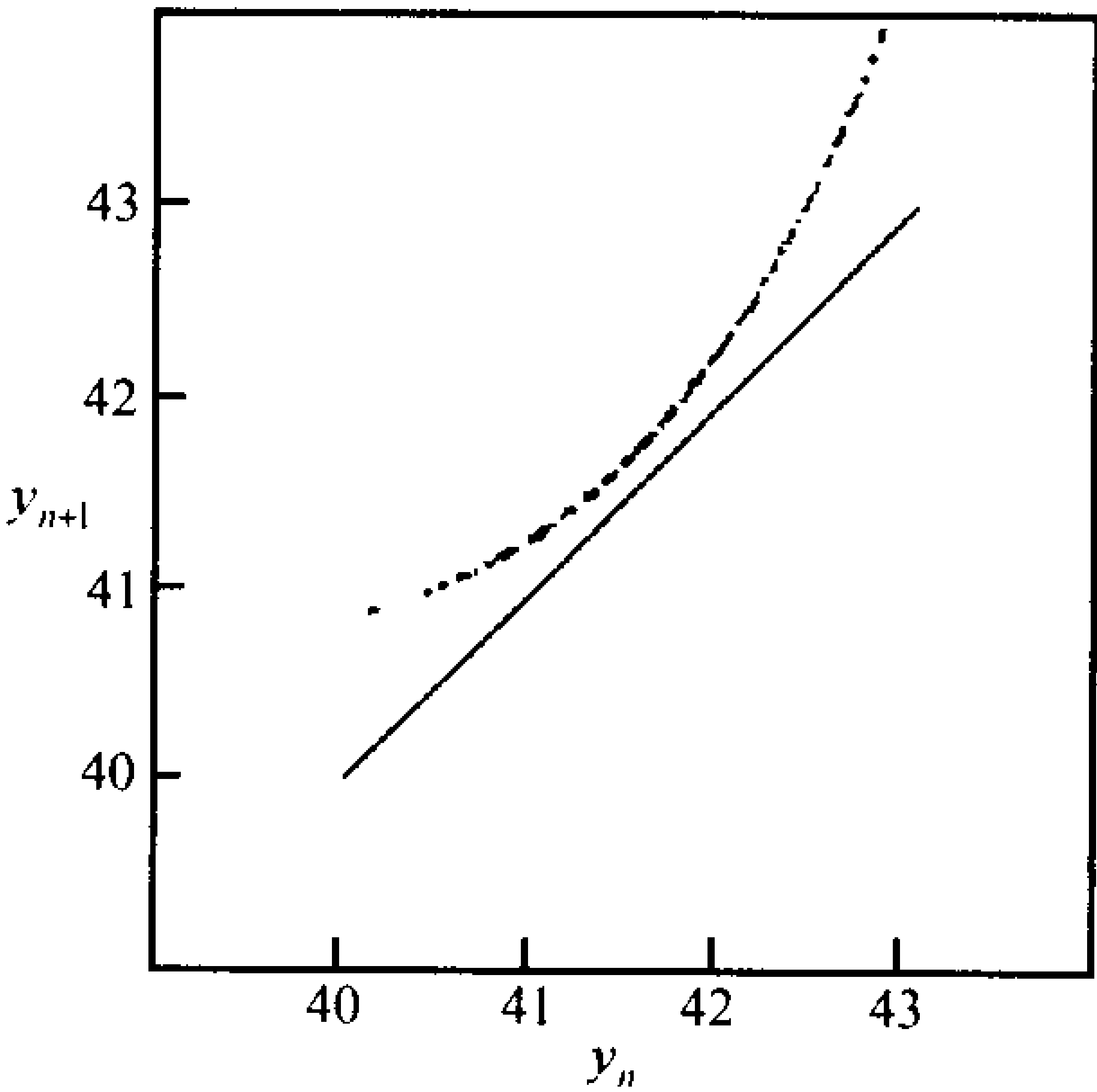


图 2.2.14　Lorenz 系统对应的 Poincaré 映射 (当 r 略大于 $r_c = 166$ 时)

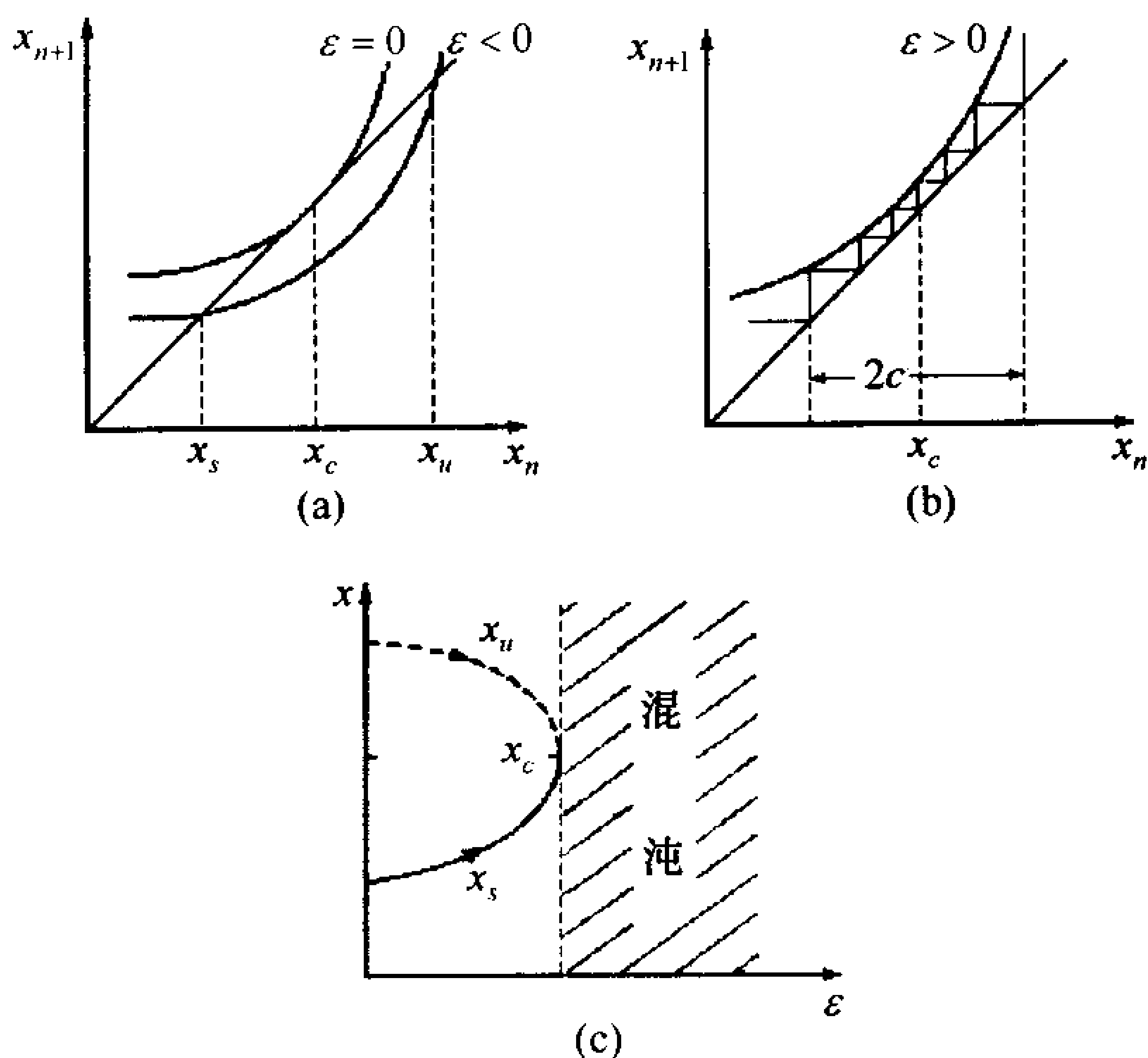


图 2.2.15　I 型间歇的机制

(a) 当 $\varepsilon=r-r_c\leqslant 0$ 时的 Poincaré 映射; (b) 当 $\varepsilon>0$ 时的 Poincaré 映射和运动轨迹, 注意不动点 "幽灵" x_c 吸引左边的轨道并把它们向右边排斥; (c) 倒切分岔

当 $r>r_c$ 映射不具有稳定不动点, 但是表现出一种对于不动点的 "记忆", 因为在 x_c 附近轨迹的运动减慢了. 为了穿越映射曲线与分角曲线之间的狭窄通道, 需要许多次迭代, 这就导致当 r 值稍超过 r_c 时出现在图 2.2.13 中的长的层流区. 当轨迹通过这个通道后, 运动变为混沌的, 直到重新进入 x_c 附近时又开始新的规则过程.

实际上从离散的抛物线映射 F_μ 的数值计算发现, 当 $\mu=1+\sqrt{8}$ 时, 此时 $\mu>\mu_\infty=3.5699\cdots$, 展现出周期 3 的循环轨道, 并且不断分岔, 即在混沌区中有一个窗口, 下面图 2.2.16 显示了当 r 值大于或小于 r_c 时的不同的迭代结果. 在稍大于 r_c 时有一个规则的周期 3 循环, 但小于 r_c 时则发生被混沌隔断的层流区.

在图 2.2.17 中画出了当 $r=r_c$ 时, $F_\mu(x)$ 的三次迭代, 由它可以解释上面这种奇怪的动力学行为. 有 3 个不动点, 它们在 $r<r_c$

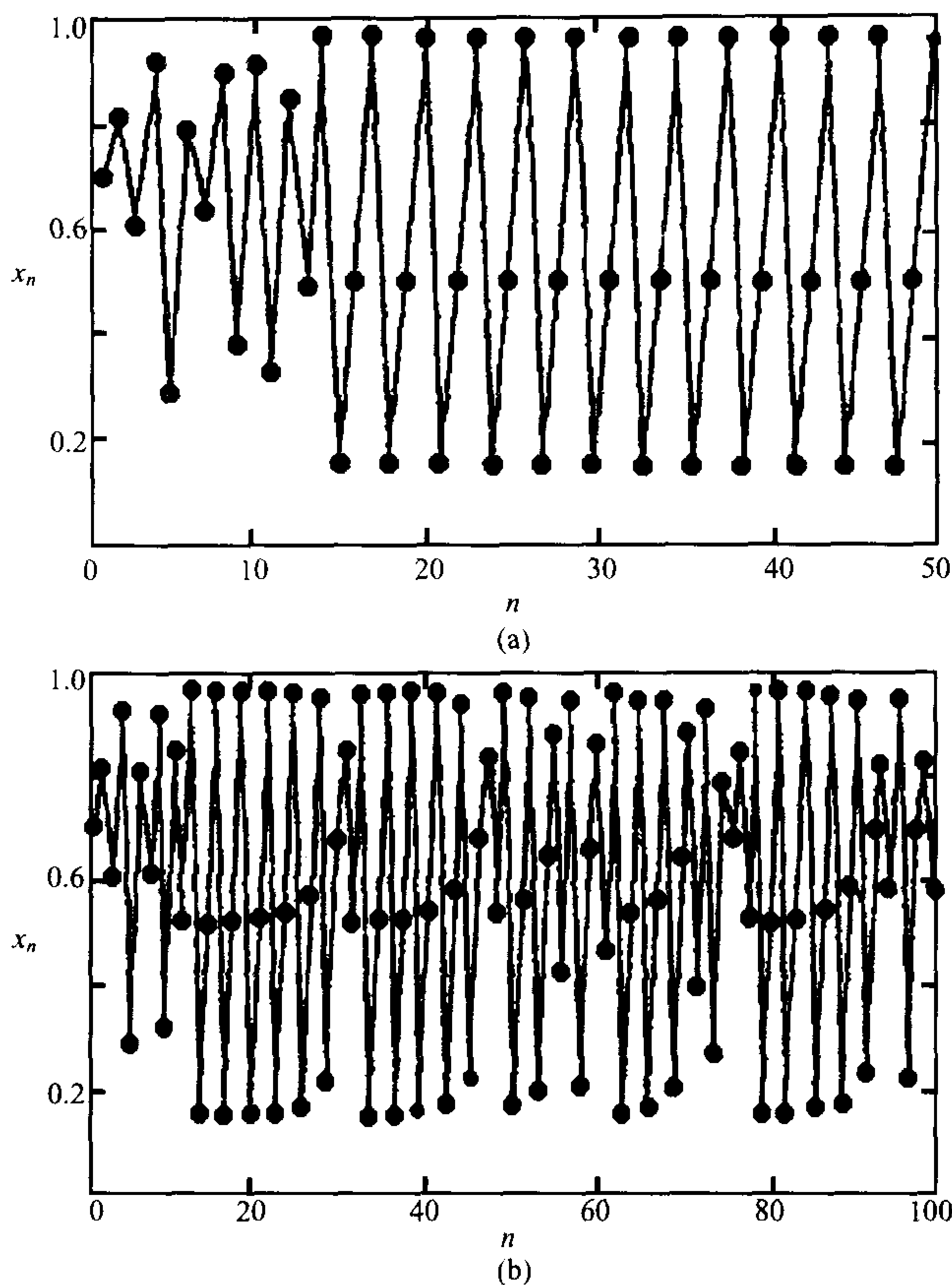


图 2.2.16　从 0.7 开始的 F_μ 的迭代结果

(a) 在稳定的 3 周期循环区，$r = r_c - 0.02$; (b) 在阵发区 $r = r_c - 0.0002$

时是不稳定的，因而导致 I 型间歇. 应该注意，倒切分岔 (不同于使不动点数目成倍增加的叉形分岔) 提供了能在抛物线映射中产生非偶数个不动点的唯一机理.

在讨论周期解及混沌形成的道路中，在前面已经涉及到倍分叉、切分叉等概念，分叉、分岔、分支实为同一涵义，对它的研究已经在动力学的讨论中占有很重要的地位. 下面通过一个例子介绍 Hopf 分支以及有关的理论结果.

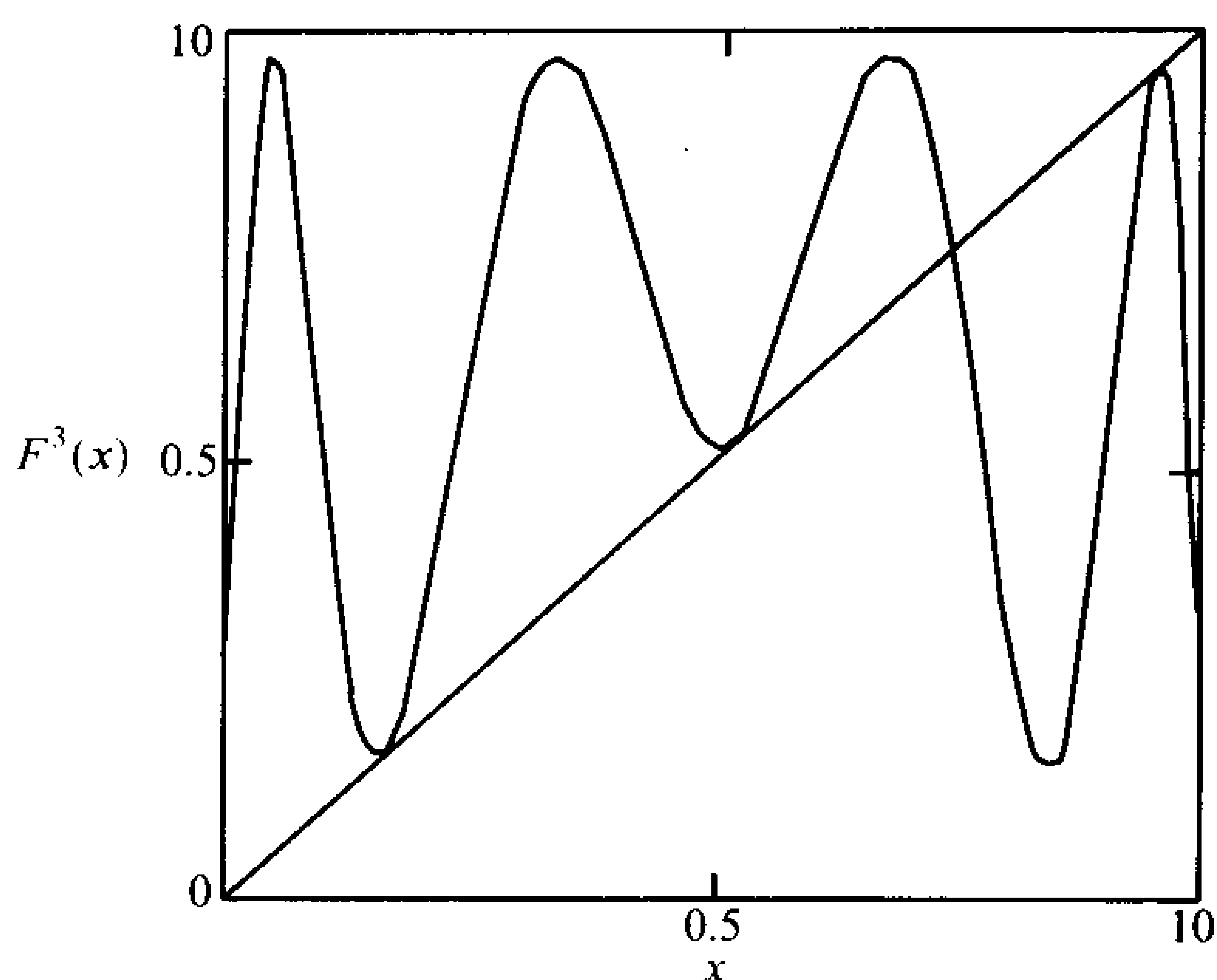


图 2.2.17　当 $r=r_c$ 时的三次迭代映射 $F_\mu^3(x)$

例 2.2.16　考虑如下的二维带参数 c 的动力学模型

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t}=x_2-x_1(x_1^2+x_2^2-c), \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t}=-x_1-x_2(x_1^2+x_2^2-c). \end{cases} \tag{2.2.61}$$

易见它的一次近似组为

$$\begin{cases} \dfrac{\mathrm{d}x_1}{\mathrm{d}t}=cx_1+x_2, \\ \dfrac{\mathrm{d}x_2}{\mathrm{d}t}=-x_1+cx_2, \end{cases} \tag{2.2.62}$$

其系数矩阵

$$A=\begin{pmatrix} c & 1 \\ -1 & c \end{pmatrix},$$

它的特征值是一对共轭复根 $c\pm\mathrm{i}$ $(\mathrm{i}=\sqrt{-1})$. 易知原点是 (2.2.61) 的平衡点. 当 $c<0$ 时, 原点是稳定的, 当 $c>0$ 时, 原点是不稳定的, 此时存在稳定的极限环 $x_1^2+x_2^2=c$ $(c>0)$. 当 $c=0$ 时, A 的两个根是一对虚根, c 在从小于 0 穿过等于 0 到大于 0 整个

参数的变化过程，不动点原点从稳定发生跳跃变成不稳定且分支成为一个稳定的极限环，这种现象称为 **Hopf 分支**.

以下介绍 Hopf 定理，现考虑一般的 n 维系统

$$\frac{\mathrm{d}x}{\mathrm{d}t} = F(x,\mu), \tag{2.2.63}$$

$x \in \mathrm{R}^n$, μ 是参数.

定理 2.2.24 假定

(1) $F(0,\mu)=0$, μ 属于包含 0 的开区间， $0 \in \mathrm{R}^n$ 是 F 的孤立零点;

(2) F 关于 x,μ 在 $(0,0)\in \mathrm{R}^n\times\mathrm{R}^1$ 的邻域内解析;

(3) 置 $A(\mu)$ 为 F 的变分矩阵，即 $A(\mu)=D_xF(0,\mu)$. 设 $A(\mu)$ 有一对共轭复根 λ 和 $\overline{\lambda}$, 而

$$\lambda(\mu)=\alpha(\mu)+\mathrm{i}\ \omega(\mu),$$

其中 $\omega(0)=\omega_0>0$, $\alpha(0)=0$, $\alpha'(0)\neq 0$;

(4) 设 $A(\mu)$ 其它的 $n-2$ 个特征值均具负实部，
则 (2.2.63) 有一簇周期解 $P_\varepsilon(t)$ 出现在 $\mu=\mu^H(\varepsilon)$, 其中 $\varepsilon\in[0,\varepsilon_H)$, 具体地说，存在 $\varepsilon_H>0$ 及解析函数

$$\mu^H(\varepsilon)=\sum_{i=2}^{\infty}\mu_i^H\varepsilon^i \quad (0<\varepsilon<\varepsilon_H),$$

$\mu^H(\varepsilon)$ 不恒为 0 而且对 $\varepsilon\in(0,\varepsilon_1)$, $\varepsilon_1\in(0,\varepsilon_H]$, $\mu^H(\varepsilon)$ 或是严格正的或是严格负的. 对每个 $L>\dfrac{2\pi}{\omega_0}$, 存在 $x=0$ 的邻域 η 以及包含 0 的开区间 J, 使得对 $\mu\in J$, (2.2.63) 有唯一的属于一簇周期解 $P_\varepsilon(t)$ 不恒为常数的具有周期 L 的周期解.

又知 $P_\varepsilon(t)$ 的周期 $T^H(\varepsilon)$ 是解析函数

$$T^H(\varepsilon)=\frac{2\pi}{\omega_0}\left[1+\sum_{i=2}^{\infty}T_i^H\varepsilon^i\right] \quad (0<\varepsilon<\varepsilon_H).$$

实际上两个 Floquet 指数当 $\varepsilon\downarrow 0$ 时趋向于 0. 一个是 0, 另一个是解析函数

$$\beta^H(\varepsilon)=\sum_{i=2}^{\infty}\beta_i^H\varepsilon^i\quad(0<\varepsilon<\varepsilon_H).$$

如果 $\beta^H(\varepsilon)<0$. 那么周期解 $P_\varepsilon(t)$ 是轨道渐近稳定的; 如果 $\beta^H(\varepsilon)>0$, 那么 $P_\varepsilon(t)$ 是不稳定的.

参 考 文 献

[1] 卢侃等编译, 混沌动力学, 上海翻译出版公司, 1990.

[2] 王东生、曹磊编著, 混沌、分形及其应用, 中国科学技术大学出版社, 1995.

[3] 陈予恕、唐云等编, 非线性动力学中的现代分析方法, 科学出版社, 1992.

[4] H. 舒斯特著, 混沌学引论, 四川教育出版社, 1994.

[5] 杨维明编著, 时空混沌和耦合映像格子, 上海科技教育出版社, 1994.

第三章　神经网络中的非线性动力学问题及分析

神经网络是一门新兴的、综合性、交叉性很强的学科. 神经网络中的非线性动力学问题是其中的一个重要组成部分，它的研究不仅有助于理解神经网络数学理论的依据与背景，而且提供了应用的基本思想及可能的途径.

本章将对各类神经网络模型中的非线性动力学问题以及分析的方法与结果作详细的介绍，这部分是本书的重点.

§3.1　神经网络与动力系统

3.1.1　作为动力系统的神经网络

对于神经网络现在有许多种描述方式与看法，例如它是一套硬件设备，是一个生物组织，是一个计算机的程序，是一个识别系统，是一种数学的算法等等. 人们正在通过生物实验、计算机模拟、理论分析以及通过许多种数学的手段来研究这种网络. 数学中的许多分支已被用于神经网络的分析，例如矩阵论、群论、泛函分析、差分方程、常微分方程和偏微分方程、组合数学、运筹学、概率论、微分几何乃至范畴理论和符号逻辑等. 本书主要介绍如何从动力系统理论的角度来研究讨论神经网络.

事实上，神经网络具有固定的权重、外部的输入、内部的状态，可以将它看成以权重、外部输入作为参数的关于内部状态的行为动力学，也可以看成是权学习的动力系统. 广泛一点说神经网络可以作为是权矩阵空间和输入输出状态空间的 Dicard 乘积下的一个动力系统.

一个神经网络包含 n 个处理单元，对于第 i 个单元，有 3 个相关的实数：网络的输入信号 s_i, 行为的状态 x_i 以及输出 y_i. 输出

是行为状态变量的函数 $y_i = g_i(x_i)$，输入是所有状态和一个 $n \times n$ 实参数矩阵 —— 权重矩阵 (w_{ij}) 的函数. 一般假定 $s_i = \sum_{j=1}^{N} w_{ij} y_j$，此处 w_{ij} 看成是从第 j 个神经元到第 i 个神经元的传送线路的强度.

为了完成对神经网络的数学描述，我们通过指定表达状态随时间改变的方式的一种规则，即一种动力学使神经网络置于一个动力系统之中.

按照动力系统类型的区分，神经网络可以分成决定性的或随机的、自治的或非自治的、离散的或连续的、无延迟的或有延迟的.

3.1.2 神经网络具有三种动力学行为

神经网络的动力学模型有许多种，每个模型有它特有的动力学行为与特征. 在本章开头首先介绍神经网络的各种动力学模型所共有的动力学行为与特征，在整体上给出一些看法与概念，以避免只见树木不见森林的错觉.

神经网络的各种动力学模型可能具有以下 3 种动力学行为：

(1) 收敛：某条轨道当时间 t 越来越大时，收敛于某个平衡点或稳定状态，一部分的轨道趋于平衡点集 (某平衡点可能是稳定的，也可能是不稳定的，即轨道将远离它)；

(2) 振荡：某条轨道渐近地趋向于一个周期的轨道，该周期轨道可能是稳定的，也可能是不稳定的；

(3) 混沌：通常比较粗略地规定其轨道在有界的范围内长期行为对初值极端敏感依赖地游荡运动.

从人工神经网络现有的多种动力学模型数值模拟及理论分析，也从人们的心理状态及部分的实际应用需要出发，第一种收敛的动力学行为研究得比较多且较深入. 人们总认为动力学行为最终应当是得到信息或者可能是产生信息，很自然希望这一信息是抓得住的形式，即稳定状态的平衡点，并且这种稳定状态的获取可以很有用，例如模式识别、组合优化、把打印文件转化为口头语言

等等. 我们已经相当习惯于把信息量作为数字或字符存储起来, 而存储的方式是通过某种离散的或连续的动力学模型的实体随时随地、恰当地把这些信息量从储藏的实体中提取出来.

但是如果从生物神经网络来分析第一种动力学行为, 恰恰是相当不可信的, 甚至将引起不少矛盾. 近年来从神经生物学的研究来看, 发现人们的大脑处理信息以及人们的大脑之所以有这么多的功能, 主要因为神经网络动力系统具有第二及第三类的动力学行为. 但是神经网络的动力系统理论与应用的研究只是对第一种动力学行为的研究比较完善, 对后两种动力学行为的研究由于其复杂性还刚开始, 许多数学工具没法应用, 总之发展很慢.

例如, 假设一个神经网络是振荡的, 输入一个初始状态以后, 该网络运行的动力学行为将最终停留在一个极限环上, 即一个周期轨道上. 那么我们怎么样才能收回含在这个环中的信息呢? 或许我们会计算极限环的周期、大小、它的某个分量的 Fourier 系数的一些函数, 或者沿着环的一些函数的平均. 但事实上在环和平衡点之间有基本概念上的差异, 在一个有限维的状态空间中, 任何状态都是一些动力系统的一个平衡点, 也就是说状态空间中的可能的平衡点集是有限维的, 但是可能的环的集合是一个相当丰富的无限维空间. 平衡点是一个有限的、数学上的静态目标, 仅仅是从一个坐标系变换到另一个坐标系, 是数学上的一个常规的方法, 而这一性质对于研究工作是很理想的. 但对环的说明要依赖于特定的坐标, 在某一坐标系下一个周期轨道会相当简单, 可能只是简单的调和运动, 而同样的轨道在另外的坐标系下可能表现得非常复杂, 而这些坐标是很自然地适应网络的. 环的复杂性可能精确地反映了以数学为工具的现实系统的模型的行为, 另一方面, 这种复杂性可能只是由于我们对状态变量选取不当而造成的.

在生物神经网络中还经常发生如下的情形: 如果一个网络的输出作为另一个网络的输入, 一个有趣的问题就产生了, 若第一个网络是收敛的, 则没有问题; 如果它是振荡的, 则第二个网络就有了振荡输入, 具有振荡输入的网络的研究目前还处于初级阶段.

对于收敛的神经网络虽然尚有一些没有解决的和有趣而未提出的问题，但是已经有许多被广泛接受和公认的方法. 如何阅读网络的输出，一个网络如何作为联想记忆器、模式分析器、优化器等等来进行运作，用什么来监控以及不监控学习方法等问题已广泛达成共识.

S. Grossberg 和 S. Elias 研究了短期记忆的网络模型中的振荡动力学行为， W. Baird, Z. Li 和 J. Hopfield 研究了作为稳定振荡的储存数据的网络和训练权重的方法连同嗅觉球管模型.

对于混沌神经网络，存在着许多令人困惑的问题，对这些问题的研究虽有一定的进展，但离完美的解决还相差很远. 例如一个混沌轨道的极限集一般呈分形结构，那么在什么意义下它代表信息？我们如何从分形中提取信息？如何使用混沌的输出作为另一个网络的输入？如何训练具有混沌动力学行为的神经网络的权重？在什么情况下一个混沌的神经网络动力系统是鲁棒的？如何确定一个神经网络是的确具有混沌的？

虽然困惑如此之多，但是有许多好的动因促使我们必须去理解及利用振荡的和混沌的神经网络. 在对大脑的动力学实验中发现了一个令人震惊的事实：脑动力学行为从来都不是收敛的，一般都是振荡的或是混沌的. 除了生物的原因外还有工程上的原因，在工程方面，其动力学行为更为丰富. 但是我们还不知道这些丰富的动力学行为有什么应用.

生物学上又提示我们，神经网络构成的大脑尽管不是收敛的，却是高度重现的. 工程上的答案是对于同样数目的单元，收敛重现网络动力学行为的所有表现比前馈式网络更丰富. 如前馈式网络不能用于竞争学习，也不能用于适应性回应，但实验表明重现网络可以学习得很快，有的文章提出了可以完成相当复杂的阶段性任务的一般重现网络不一定收敛的学习算法.

本章仅限于较详细地介绍一些已经成熟的与公认的动力学分析的结果，对一些有争议的、目前尚不成熟的内容只能简单地提及.

3.1.3 神经网络动力系统研究的意义和作用

从神经网络的研究与发展的历史来看，无论是生物神经网络还是人工神经网络，当对于它们的动力学模型以及其中的非线性动力学问题研究获得新的进展时，这些神经网络系统的性质、功能及应用也就有了强有力的理论支撑，也就能被广泛引用及采用.

20 世纪 80 年代初期，美国加州技术学院的优秀生物物理学家 John J. Hopfield 博士发表了一篇十分重要的论文，文中动力学模型被用于全联接的神经网络以后称为 Hopfield 神经网络的描述，并用集成电路硬件实现了这个动力学的系统，又通过广义能量及 Lasalle 不变性原理给出了十分重要的动力学分析的结果：神经网络的状态即动力学模型中的流最终收敛于平衡点集. 这就为联想记忆及优化的性能与功效提供了强有力的理论基础，又为实际的应用提供了可靠的依据，从而使人工神经网络的研究进入了一个新的高潮时期.

与 Hopfield 几乎同时， PDP 并行分布处理研究小组以 David E. Rumelhart 以及 James L. Mc Chelland 为首的十多人，他们对从 R^n 到 R^m 的高维非线性映射用前馈网络，简称为 BP 网络的权的优化选择实现了映射的逼近. 这个理论上的突破使神经网络的研究与应用从此又掀起了新高潮.

这个时期的一个著名的神经生物实验与模型的计算机模拟，即当电位和电流在相当长度的轴突上保持均匀的空间箝制条件时，用周期性脉冲刺激乌贼的区轴突的实验，显示了神经元的复杂动力学行为，不仅可能吸引到不动点，还可能吸引到极限环及混沌吸引子. 这个重要的实验进一步揭示了神经元的非线性动力学的复杂行为，不久就引发了 20 世纪 90 年代初的 Aihara 混沌神经元网络模型的建立，并使其在动态的联想记忆等方面获得了应用，从而把神经网络研究又推向了更新的高潮.

生物的神经系统是一个由大量的非线性元件通过广泛联结所构成的多级系统，正是由于各级层次上自系统的非线性，使其上一层次具有下一层次自系统所没有的一些“突现性质”(emergent

property), 从而使神经系统的功能表现出极大的复杂性. 可以毫不夸张地说, 没有非线性也就没有生命, 也正因为如此, 在神经系统的各个层次上都有着大量的非线性动力学问题, 只是由于研究上的困难, 使得对神经系统各个层次的非线性动力学研究推迟到 20 世纪 80 年代才开始.

人工神经网络系统是采用物理可实现的系统来模仿人脑神经细胞的结构和功能的系统, 如果这种模仿做得好, 它必然有生物神经网络系统所具有的复杂动力学性质. 人工神经网络模型对于神经生物学性质模拟得越好, 就越能体现出复杂的动力学行为与性质, 即对人脑功能的模仿效果越来越好. 一般地说, 人工神经网络的动力学研究经常是紧跟着生物神经网络的动力学研究而进行的. 在这一章中, 我们并不把人工的与生物的神经网络模型中的动力学问题分开来讨论, 而是放在一起讨论, 用动力学的观点统一处理这两类模型, 这样有利于还其本来的真面目, 并为神经网络的系统研究与开发提供一条比较自然的、效果比较显著的途径.

§3.2 单个神经元的动力学模型与问题

单个神经元的动力学模型最早有 1943 年的 M–P 模型 (由 McCulloch 和 Pitts 提出), 后来有 1961 年的 Caianiello 模型及 1971 年的 Nagumo–Sato 模型, 到 1990 年, Aihara 提出了混沌神经元模型, 它们揭示的动力学行为有简单的不动点、较复杂的周期的极限环, 直到更复杂的混沌奇异吸引子. 随着这些模型的提出, 吸引子的形态越来越复杂, 而这些模型的应用也越来越广泛.

3.2.1 McCulloch–Pitts 模型

1943 年 McCulloch–Pitts 提出了早期的神经元的 M–P 模型, 它的特点为输入、输出都是二值的变量, 结构是固定的且权也是固定的. 它的权分为兴奋性突触权和抑制性突触权两类. 如果兴奋性突触能激活, 那么神经元被兴奋, 输出为 1; 如果抑制性突触能激活, 那么神经元被抑制, 输出为 0. 而兴奋性突触能否激活要看它

的累加值是否大于一个阈值. 记 x_{ij} 为兴奋性突触的输入 (第 j 个神经元突触对第 i 个神经元的输入) ($j=1,2,\cdots,n_1$), $\widetilde{x}_{ik}$ 为抑制性突触的输入， θ_i 为阈值. 设 $E_i=\sum\limits_{j=1}^{n_1}x_{ij}$, $I_i=\sum\limits_{k=1}^{n_2}x_{ik}$, 则输入 x_{ij}, x_{ik} 与输出 y_i 的转换关系为

$$y_i=\begin{cases}1, & \text{当 } I_i=0,\, E_i\geqslant\theta_i \text{ 时},\\ 0, & \text{当 } I_i=0,\, E_i<\theta_i \text{ 或 } I_i>0 \text{ 时}.\end{cases}$$

如果兴奋性突触权 w_{ij} 用 +1 表示，抑制性突触权 $\widetilde{w}_{ik}$ 用 −1 表示，则 M–P 模型可以写成如下数学形式:

$$y_i=H\left(\sum_{j=1}^{n_1}w_{ij}x_{ij}+\sum_{k=1}^{n_2}\widetilde{w}_{ik}\widetilde{x}_{ik}-\theta_i\right),$$

这里 H 为 0, 1 两值的单位阶梯函数. 这个模型是对第 i 个单个神经元的 M–P 模型，它同有逻辑门组成的逻辑式的实现区别不大，现在很少有人单独使用，但它的思想直到现在还是许多神经网络研究的基础.

如果记 $x_i(t)$ 为第 i 个神经元在 t 时刻的输出值，则 M–P 模型可以写为以下形式:

$$x_i(t+1)=H\left(\sum_{j=1}^{N}w_{ij}x_i(t)-\theta_i\right), \tag{3.2.1}$$

或将 i 下标去掉可以写成最简单的形式:

$$x(t+1)=H(Wx(t)-\theta).$$

3.2.2 Caianiello 模型

如果考虑 N 个神经元对某个 (例如第 i 个) 神经元的激发的影响不仅仅对一个时间单位 (即对下一个时刻) 后，而对 $r+1$ 个时

间单位后均有影响, 这里 r 是不小于 1 的某自然数. 仍记 $x_i(t+1)$ 为第 i 个神经元在时间 $t+1$ 时的输出值 (t 取离散时间值, 即为自然数), 该输出值仍考虑取值为 1 (激发状态) 或 0 (非激发状态). 取 H 为单位跳跃函数, 即

$$H(u)=\begin{cases}1, & u\geqslant 0,\\ 0, & u<0.\end{cases}$$

由于实际的生物神经元具有不应性, 即在神经元激发后的某些时刻, 对自身的输出保留影响, 设立即发出反应使动作电位产生并使神经元的输出产生变化. 记 $w_{ij}^{(r)}$ 为第 j 个 $(j\neq i)$ 神经元的激发经 $r+1$ 个离散时间单位后, 对第 i 个神经元的影响的连接权值. $w_{ii}^{(r)}$ 为第 i 个神经元的激发经 $r+1$ 个离散时间单位对自身保留的影响, 也就是 $t-r$ 时刻第 j 个神经元的输出值在 $t+1$ 时刻对自身输出的影响的强度系数. θ_i 是第 j 个 $(j\neq i)$ 神经元的 "全或无" 的阈值.

Caianiello 建立的数学模型为

$$x_i(t+1)=H\left(\sum_{j=1}^{N}\sum_{r=0}^{t}w_{ij}^{r}x_j(t-r)-\theta_i\right), \tag{3.2.2}$$

也可简单地记为

$$x(t+1)=H\left(\sum_{r=0}^{t}w^{r}x(t-r)-\theta\right).$$

3.2.3 Nagumo–Sato 神经元模型

1971 年日本学者 Nagumo 和 Sato 在 Caianiello 模型的基础上作了如下的一些假设:

(i) $A(t)$ 为离散时刻 t 的输入强度;

(ii) $w_{ij}^{r}=-\alpha k^{r}$, α 为一正参数, k 为不应性衰变率, $0\leqslant k\leqslant 1$, 即假设过去某时刻的激发导致的不应性影响随着时间 r 指数地减少.

他们建立的神经元模型的数学方程式为

$$x(t+1) = H\left(A(t) - \alpha\sum_{r=0}^{t} k^r x(t-r) - \theta\right). \qquad (3.2.3)$$

现在定义神经元的内部状态 $y(t+1)$ 为

$$y(t+1) = A(t) - \alpha\sum_{r=0}^{t} k^r x(t-r) - \theta. \qquad (3.2.4)$$

以下建立关于 $y(t)$ 与 $y(t+1)$ 的动力学关系方程，注意到

$$x(t+1) = H(y(t+1)),$$

也就是说由 $y(t+1)$ 的动力学行为可以推知 $x(t+1)$ 的动力学行为．记

$$a(t) = A(t) - kA(t-1) + (k-1)\theta, \qquad (3.2.5)$$

又从 (3.2.4) 知

$$y(t) = A(t-1) - \alpha\sum_{r=0}^{t-1} k^r x(t-r-1) - \theta. \qquad (3.2.6)$$

以下证明

$$y(t+1) = ky(t) - \alpha x(t) + a(t), \qquad (3.2.7)$$

考虑到 $x(t) = H(y(t))$, 即有

$$y(t+1) = ky(t) - \alpha H(y(t)) + a(t). \qquad (3.2.8)$$

在 (3.2.6) 两边乘 k 得

$$\begin{aligned} ky(t) &= k(A(t-1) - \theta) - \alpha\sum_{r=0}^{t-1} k^{r+1} x(t-r-1) \\ &= k(A(t-1) - \theta) - \alpha\sum_{s=1}^{t} k^s x(t-s) \end{aligned}$$

$$
\begin{aligned}
&= k(A(t-1)-\theta) - \alpha\sum_{r=0}^{t} k^r x(t-r) + \alpha x(t) \\
&= A(t) - \theta - a(t) - \alpha\sum_{r=0}^{t} k^r x(t-r) + \alpha x(t) \\
&= y(t+1) - a(t) + \alpha x(t),
\end{aligned}
$$

即证得 (3.2.7). 由 $a(t)$ 的表达式 (3.2.5) 可见它为分支参数. 下面分析分支参数 $a(t)$ 改变时的神经元的内部状态 $y(t)$ 的变化, 在图 3.2.1 中显示了 $y(t)$ 随 $a(t)$ 变化的分支图 (本小节的南云 - 佐滕模型的有关图都取 $k = 0.7$).

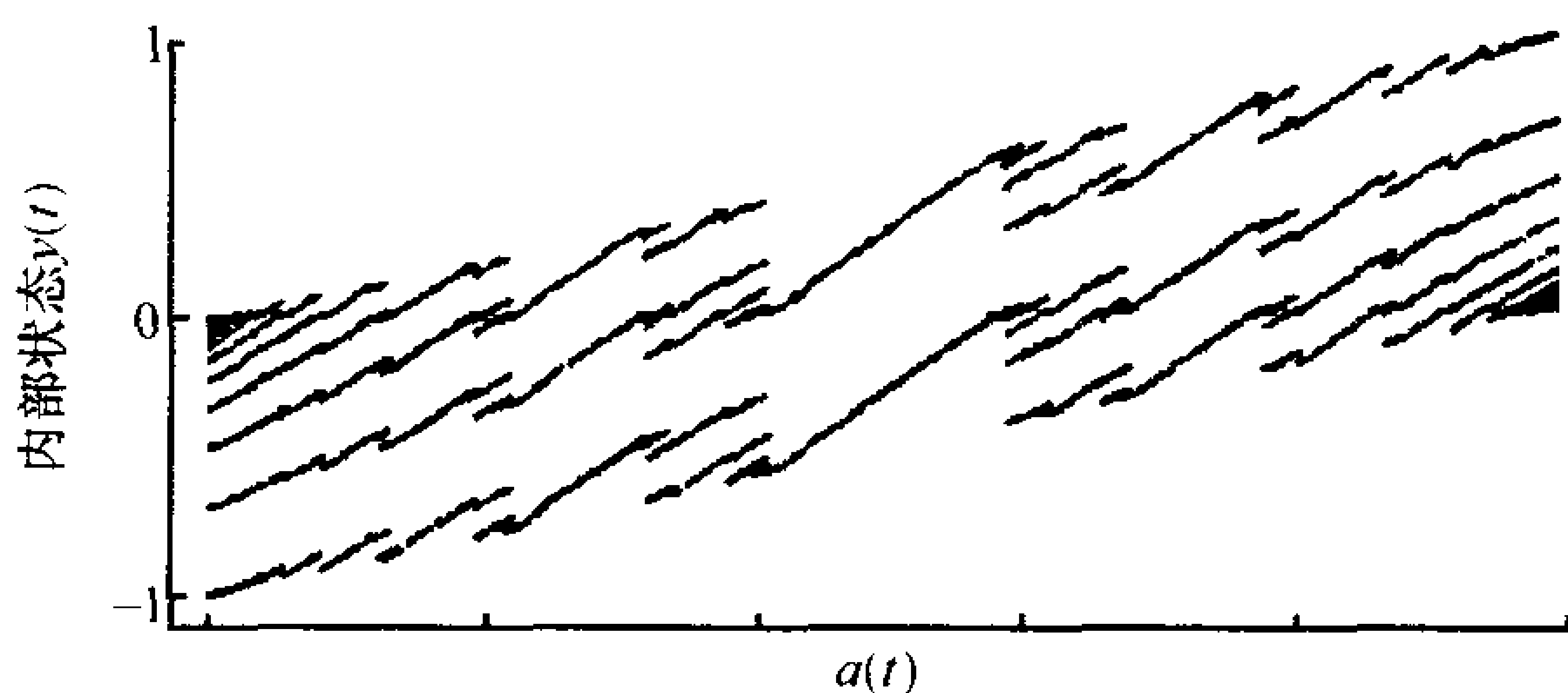


图 3.2.1　内部状态 $y(t)$ 随 $a(t)$ 变化的分支图

定义 **平均激发率** (或称为 **平均兴奋数**、**平均兴奋率**)

$$
\rho = \lim_{n\to+\infty} \frac{1}{n}\sum_{t=0}^{n-1} x(t), \tag{3.2.9}
$$

在图 3.2.2 中显示了平均兴奋数 ρ 随分支参数 $a(t)$ 变化的细小的楼梯形状的响应特性, 构成了完整的 "魔鬼楼梯".

近年来, 在研究非线性振动、非线性混沌电路中得到了有趣的具有维数自相似结构的楼梯, 如果把这个楼梯的一部分放大, 则又会出现类似的楼梯, 人们形象地称之为魔鬼楼梯. 它在数学上可建立对应关系对应于 [0,1] 区间内所有有理数的无限个台阶, 它是处处不可导且单调连续递增的 Cantor 函数. 早在 17 世纪中期, 德

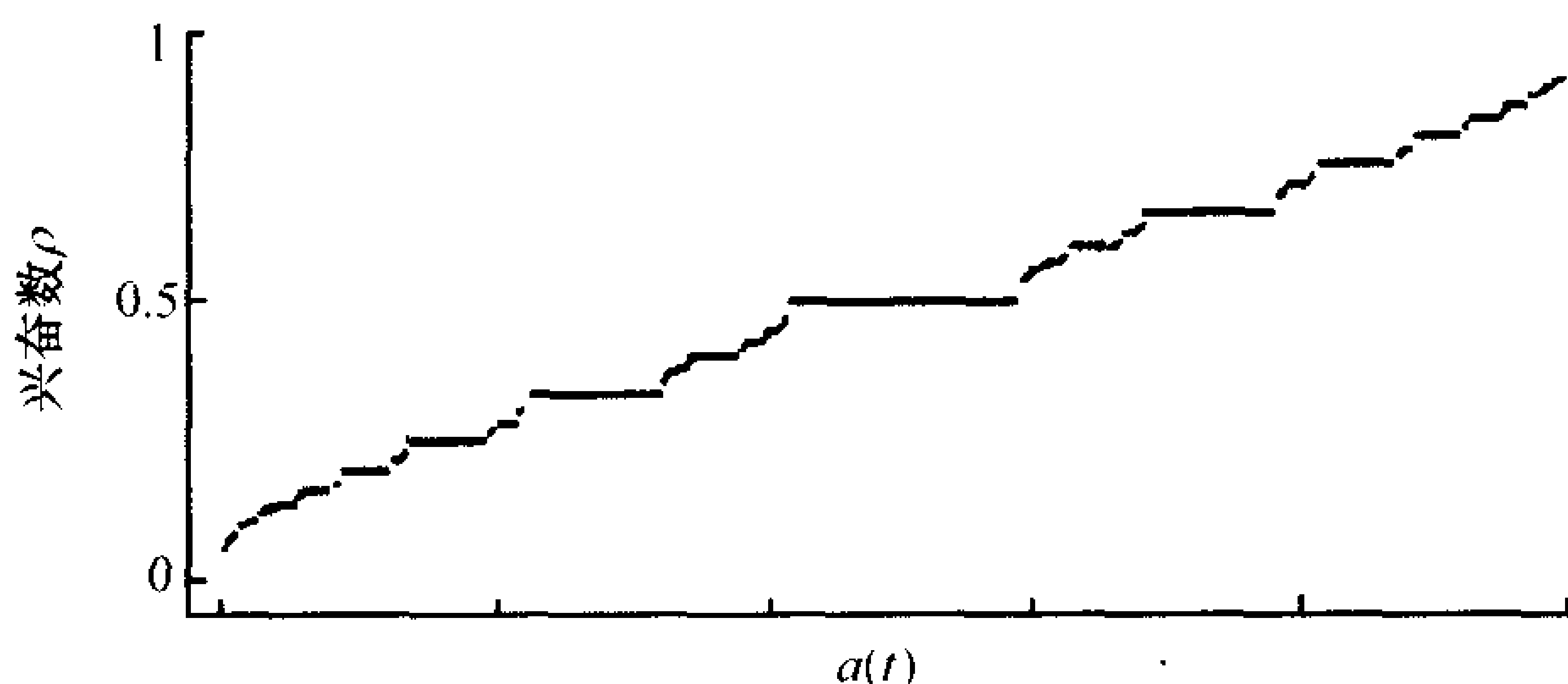


图 3.2.2 兴奋数 ρ 随 $a(t)$ 变化的响应特性图

国物理学家 Christian Huygens 对于两个背靠背挂在墙上的摆钟进行观察, 发现它们的运动会趋向于同步. 这种现象称为锁相 (锁频或共振), 它通常是在具有两个竞争步频的动力学系统中出现的. 例如一个荡秋千的小孩就是一个振荡器, 在旁看管的大人就是一个驱动者. 当一个振荡器和它的驱动者之间的相互作用充分强时, 振荡器将在无数的驱动步频率处共振. 如果将振荡器的频率与外力驱动的步频之间的关系作图, 就会得到上面所述的魔鬼楼梯.

3.2.4 一个乌贼的巨轴突上生物神经元混沌响应的实验

Nagumo–Sato 神经元模型数值模拟的结果 —— 令魔鬼楼梯的响应特性最好的检验是神经生物的实验. 以下介绍一个当电势和电流在相当长度的轴突上保持均匀的空间箝制条件时, 用周期性脉冲刺激乌贼的巨轴突, 在乌贼的神经细胞中存在着很大的轴突, 其直径可达 400~900μm. 而乌贼巨轴突的周围有一些含盐分的外液, 随着外液中盐分的增减, 在没有任何外部刺激的情况下, 乌贼巨轴突的膜电位会自发地产生周期自激振荡. 一般地活动电位以 5ms 的间隔规则地产生, 这就使我们可以将乌贼巨轴突视为神经振荡器. 如果在外部施加正弦波的周期电流刺激, 那么就可以通过外部的驱动者去调试上述神经振荡器的响应特性.

通过频闪观察器可以清楚地观察到, 当外部的输入电流正弦波的振幅和周期变化时, 神经振荡器的响应也发生改变. 由实验可

以得到下面一些结果:

(1) 乌贼巨轴突有周期性响应

例如图 3.2.3 中的 4 幅图就显示了乌贼巨轴突的这种周期性响应. 图 3.2.3(a) 是在输入正弦波的 180° 这个特定相位上采样神经膜电位的响应波形的电位 $V_1, V_2, \cdots, V_k, \cdots$ 及响应的电位的变化率值 $\dot{V}_1\left(=\dfrac{\mathrm{d}V_1}{\mathrm{d}t}\right), \dot{V}_2\left(=\dfrac{\mathrm{d}V_2}{\mathrm{d}t}\right), \cdots, \dot{V}_k\left(=\dfrac{\mathrm{d}V_k}{\mathrm{d}t}\right), \cdots$.

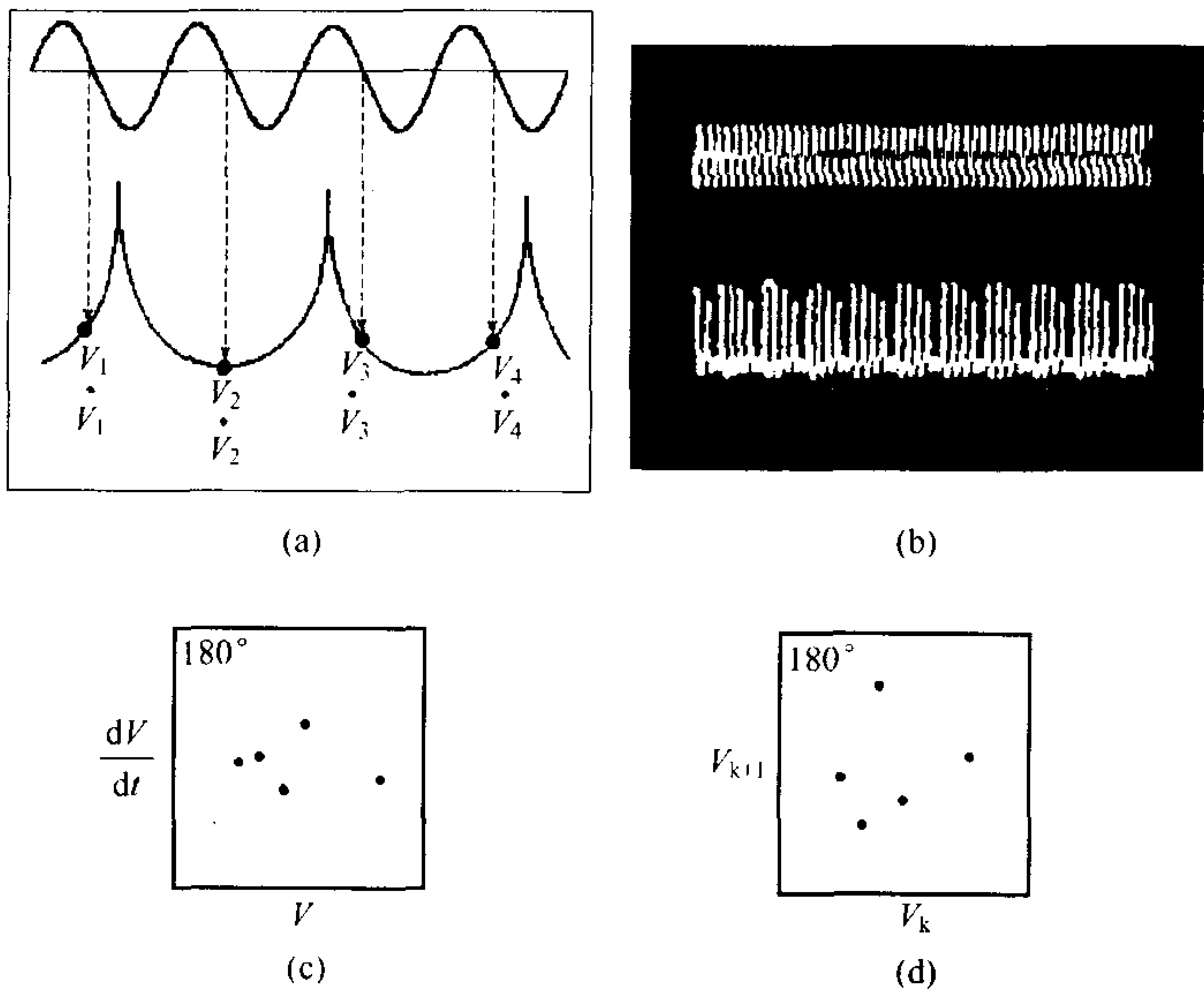


图 3.2.3

(a) 在输入电流正弦波 180° 处采样; (b) 输入波形 (上) 和响应波形 (下); (c) 状态空间 $V \times \dfrac{\mathrm{d}V}{\mathrm{d}t}$ 内轨迹的横切面; (d) 相邻两个相位对应的 V_k 与 V_{k+1} 的响应图

图 3.2.3(b) 的上图为电流的输入波形, 下图为活动电位的响应波形. 从下图中可以看出, 活动电位每产生 4 个脉冲都要有一个休止期, 并与输入保持同步的周期振荡, 这种模式的周期振荡称为 4 对 5 响应, 其平均兴奋率为 4/5. 一般地, 当基本周期是输入周期的 m 倍, 并且在基本周期内产生 n 个脉冲, 则称为 n 对 m 响

应 (n/m 响应)．把输入的周期设定于自激振荡周期的 m/n 倍附近, 就有可能产生噪声极少的状态. 如果调节输入波形的振幅和周期, 则可以产生各种各样的有理数 n/m, 从而可以预想到发生平均兴奋率为 n/m 的 n 对 m 响应. 在乌贼巨轴突实验中, m 约为 10 左右.

图 3.2.3(c) 所示为状态空间 $V \times \dfrac{\mathrm{d}V}{\mathrm{d}t}$ 内轨迹每隔一个周期 180° 的切面图.

图 3.2.3(d) 所示为输入正弦波 180° 相位时的膜电位 V_k 与下一个周期 $2 \times 180°$ 相位时的膜电位 V_{k+1} 的响应图.

在图 3.2.3(c) 与图 3.2.3(d) 中所示 5 个点对应着神经振荡器的响应周期是其输入的正弦波周期的 5 倍.

(2) 乌贼巨轴突存在准周期响应

图 3.2.4 显示了准周期的响应, 在其中自激振荡周期与外部输入周期共同存在, 在状态空间中, 其轨迹是由一个二维环面的表面卷曲而成. 图 3.2.4(a) 显示了输入波形 (上) 与响应波形 (下), 图 3.2.4(b) 显示了每隔 30° 的输入电流相位处, 相空间 $V \times \dfrac{\mathrm{d}V}{\mathrm{d}t}$ 的轨迹在横切面上的图, 它与环的横截面相对应, 渐近地描绘出一个封闭曲线.

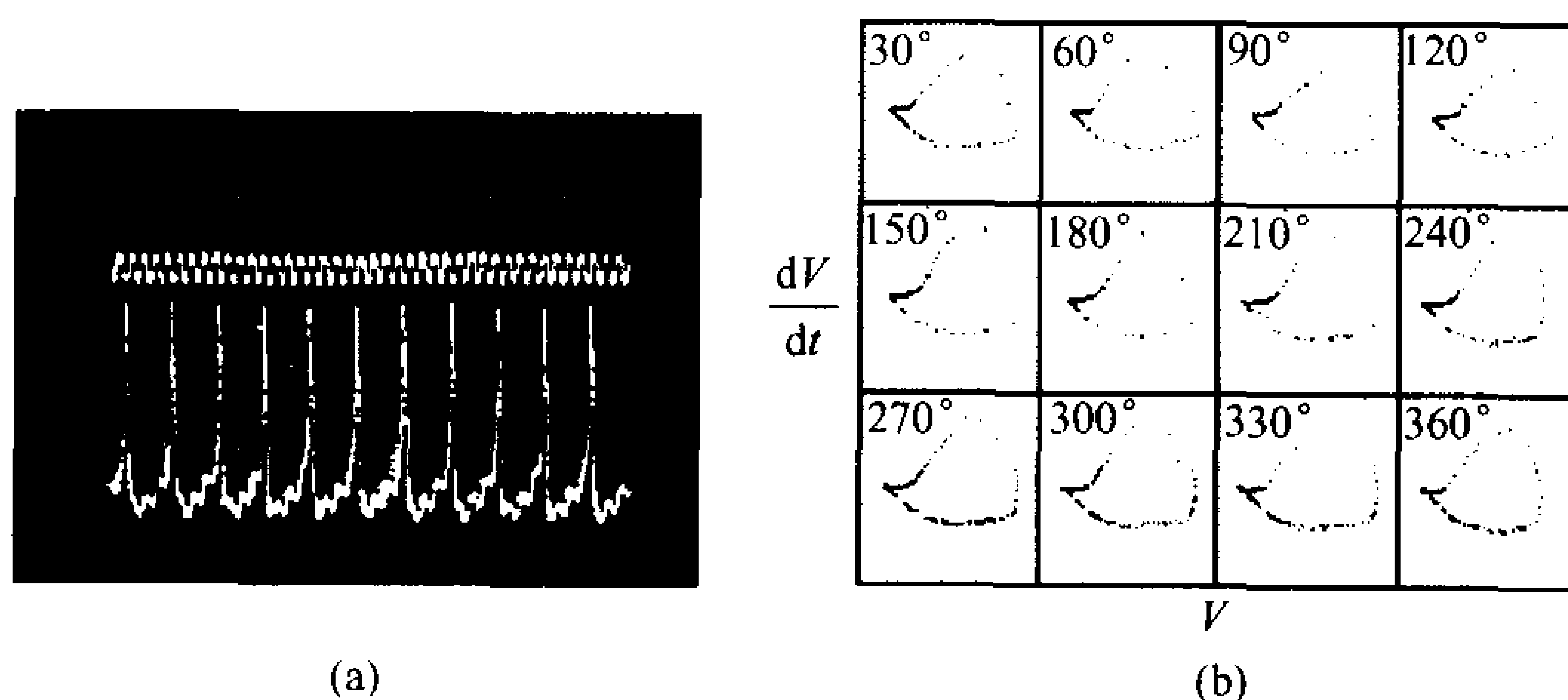


图 3.2.4　乌贼巨轴突的准周期响应

(a) 输入波形 (上) 与响应波形 (下); (b) 状态空间 $V \times \dfrac{\mathrm{d}V}{\mathrm{d}t}$ 的轨迹在每隔 30° 的横切面图

(3) 乌贼巨轴突存在混沌响应

图 3.2.5 显示了混沌响应. 图 3.2.5(b) 显示了状态空间 $V\times\dfrac{\mathrm{d}V}{\mathrm{d}t}$ 的轨迹呈现出具有混沌奇异吸引子的图案，它的分数维在 2 到 3 之间，而且具有奇异吸引子的一般特征，在局部不稳定中，整体在连续地不断徘徊，整体稳定与局部不稳定并存. 其动力学过程由图 3.2.5(b) 定性地表示出来，相位在 30° 至 120° 之间时 $V\times\dfrac{\mathrm{d}V}{\mathrm{d}t}$ 的轨迹的横截面曲线在逐渐延长，在 120° 至 180° 之间时，则开始迂回弯曲，至 180° 时完成迂回构成环状，在 180° 至 360° 之间又开始逐渐退缩，至 360° 时已退缩为很短的轨迹.

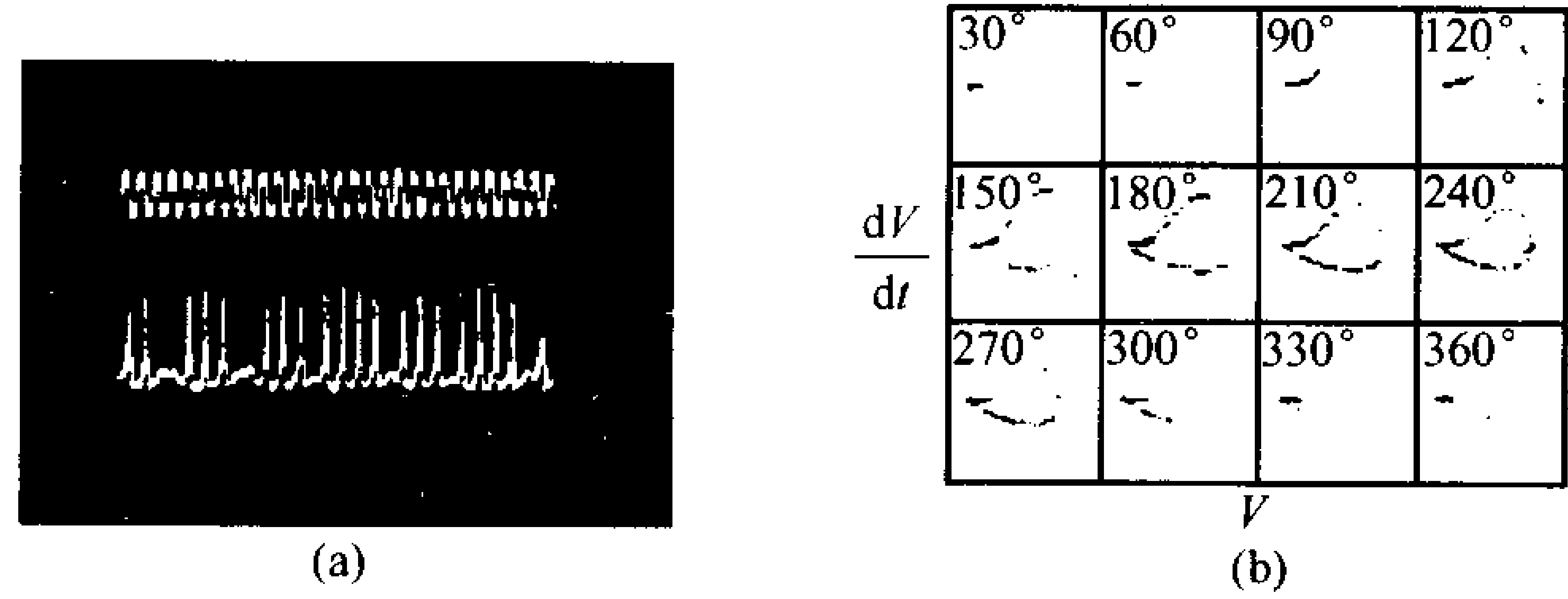


图 3.2.5 乌贼巨轴突的混沌响应

(a) 输入波形 (上) 与响应波形 (下); (b) 状态空间 $V\times\dfrac{\mathrm{d}V}{\mathrm{d}t}$ 的轨迹在每隔 30° 的横切面图

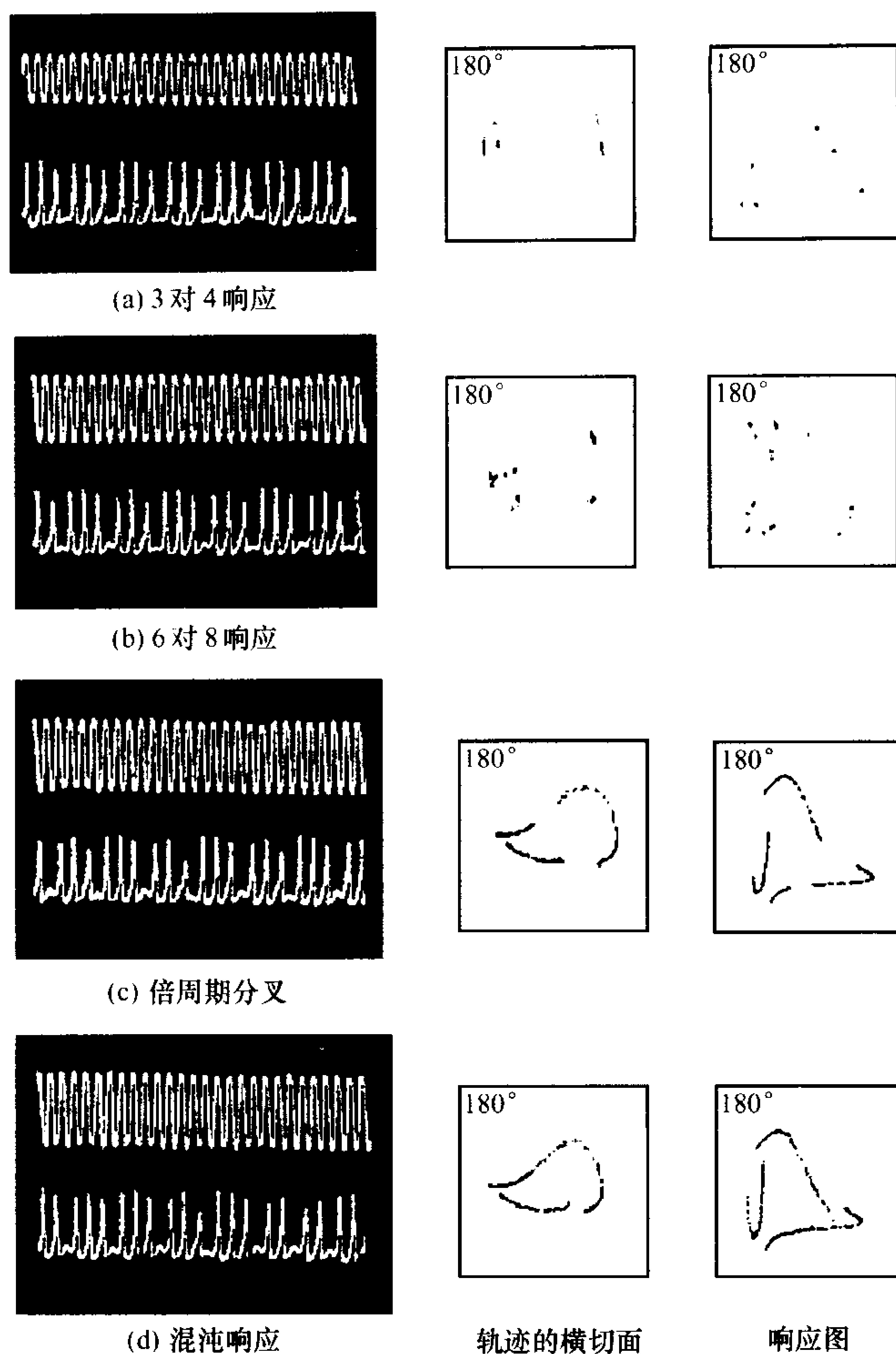
(4) 存在乌贼巨轴突从 n/m 响应到混沌响应过渡的倍周期分叉道路

图 3.2.6 所示的是从 n/m 响应到混沌响应的分支过程，可清楚地看到从 3 对 4 响应经过倍周期分叉道路向混沌过渡的情形.

(5) 存在乌贼巨轴突间歇混乱的混沌响应

从图 3.2.7 可见， 4 对 5 的响应是不稳定的，从图中轨迹的横切面来看，点密度高的部分对应于不稳定之前 4 对 5 响应的轨迹.

(6) 存在乌贼巨轴突的从准周期响应向混沌响应过渡的环面崩溃道路



(a) 3 对 4 响应

(b) 6 对 8 响应

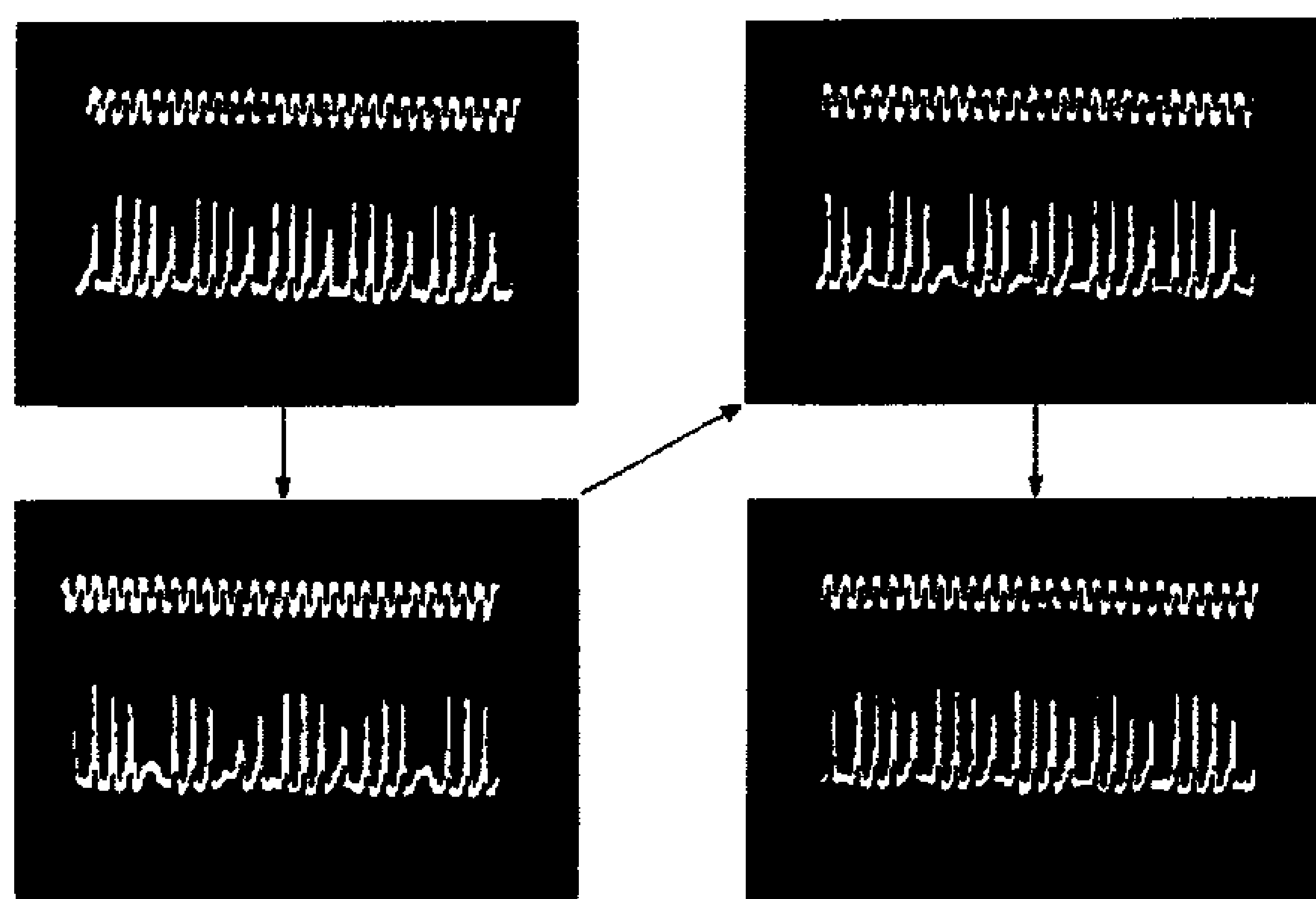
(c) 倍周期分叉

(d) 混沌响应　　轨迹的横切面　　响应图

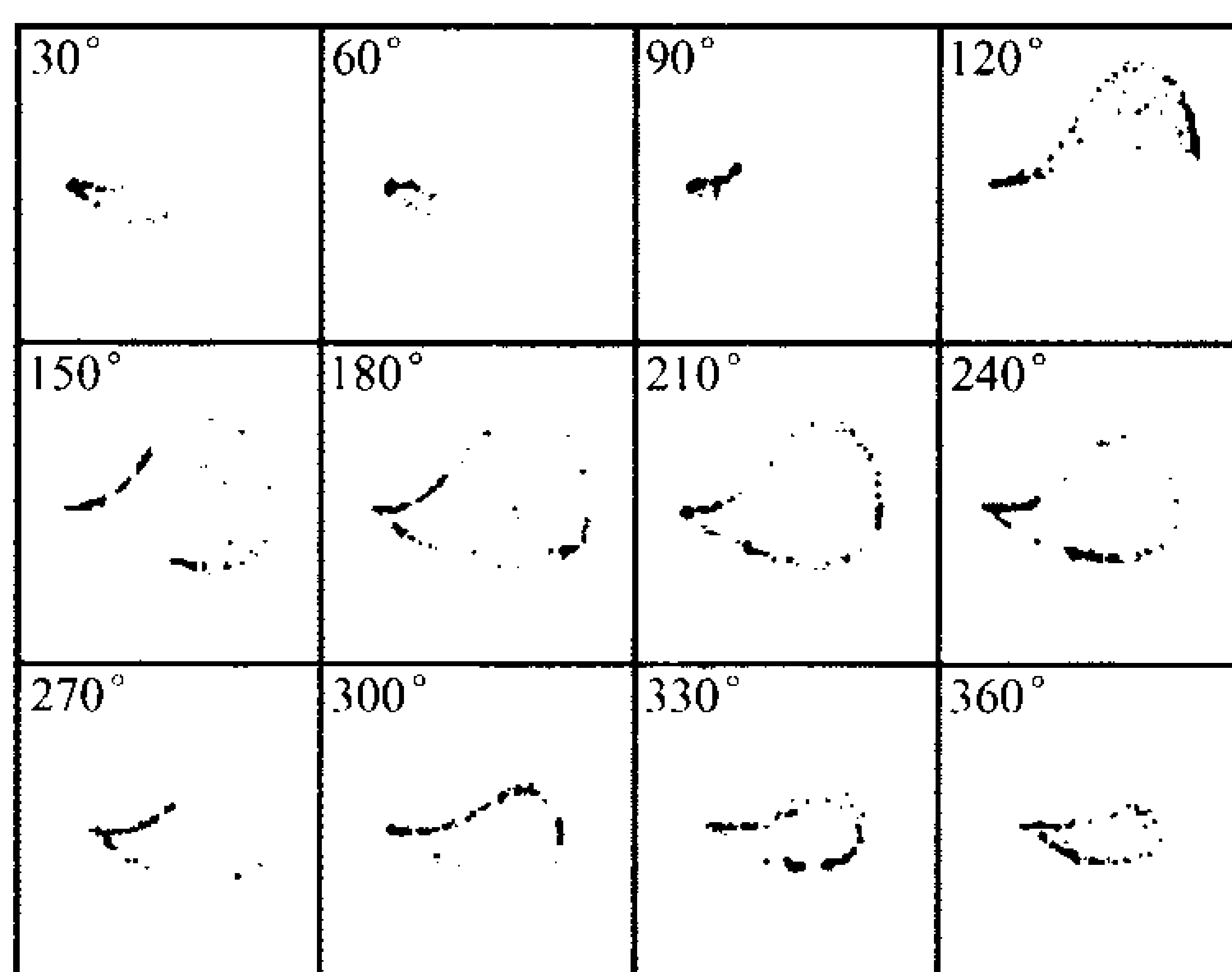
图 3.2.6　乌贼巨轴突的倍周期分叉

从图 3.2.8 可见, 图 3.2.8(a) 对应于准周期响应的二维环面, 随着输入的电流振幅的增大, 产生了变形显示于图 3.2.8(b), 并逐渐向混沌响应过渡, 显示于图 3.2.8(c).

通过以上 6 条实验结果表明乌贼自己是一个完全实在的非线性系统, 既存在着周期响应, 准周期响应, 也存在着混沌响应, 而



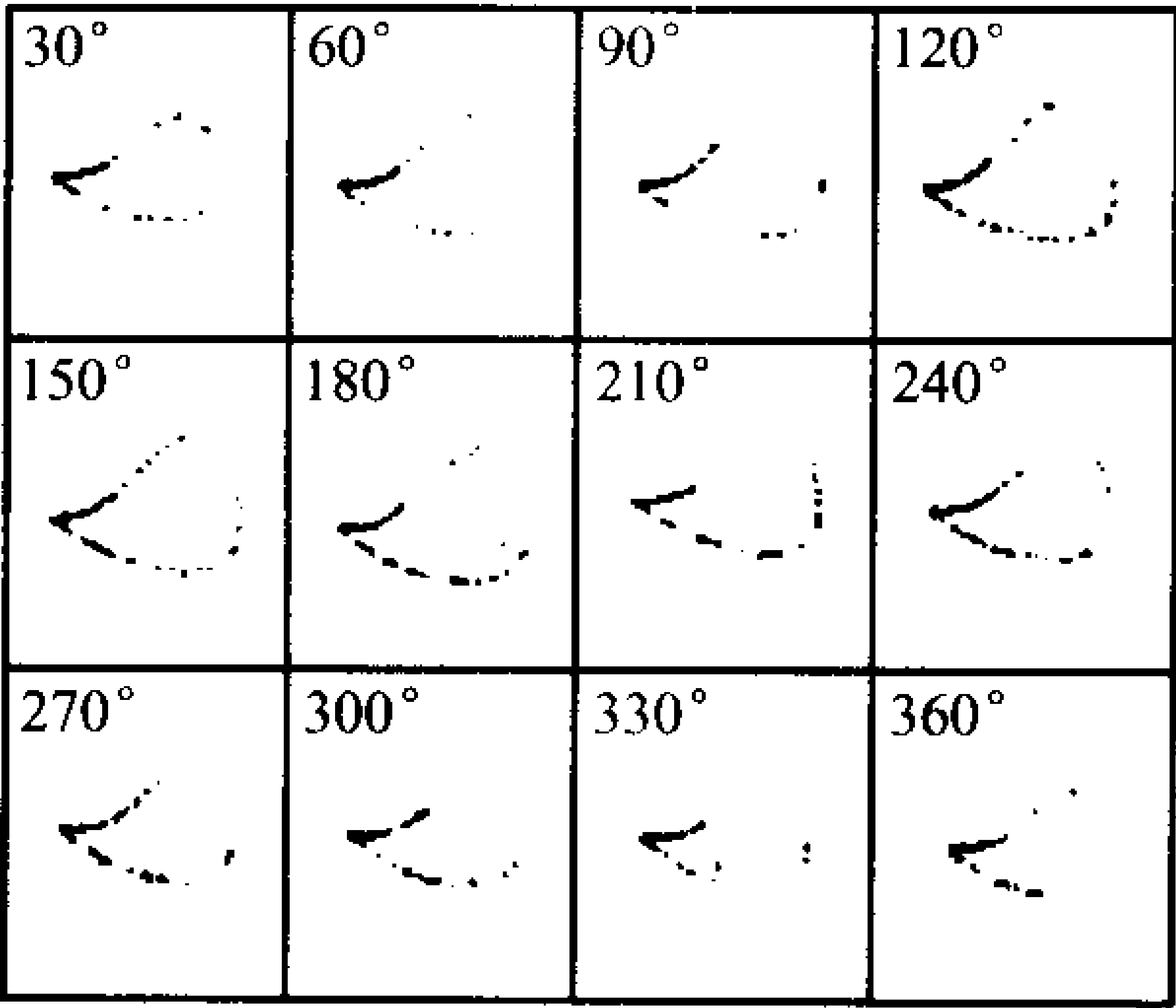
(a) 波形的变化



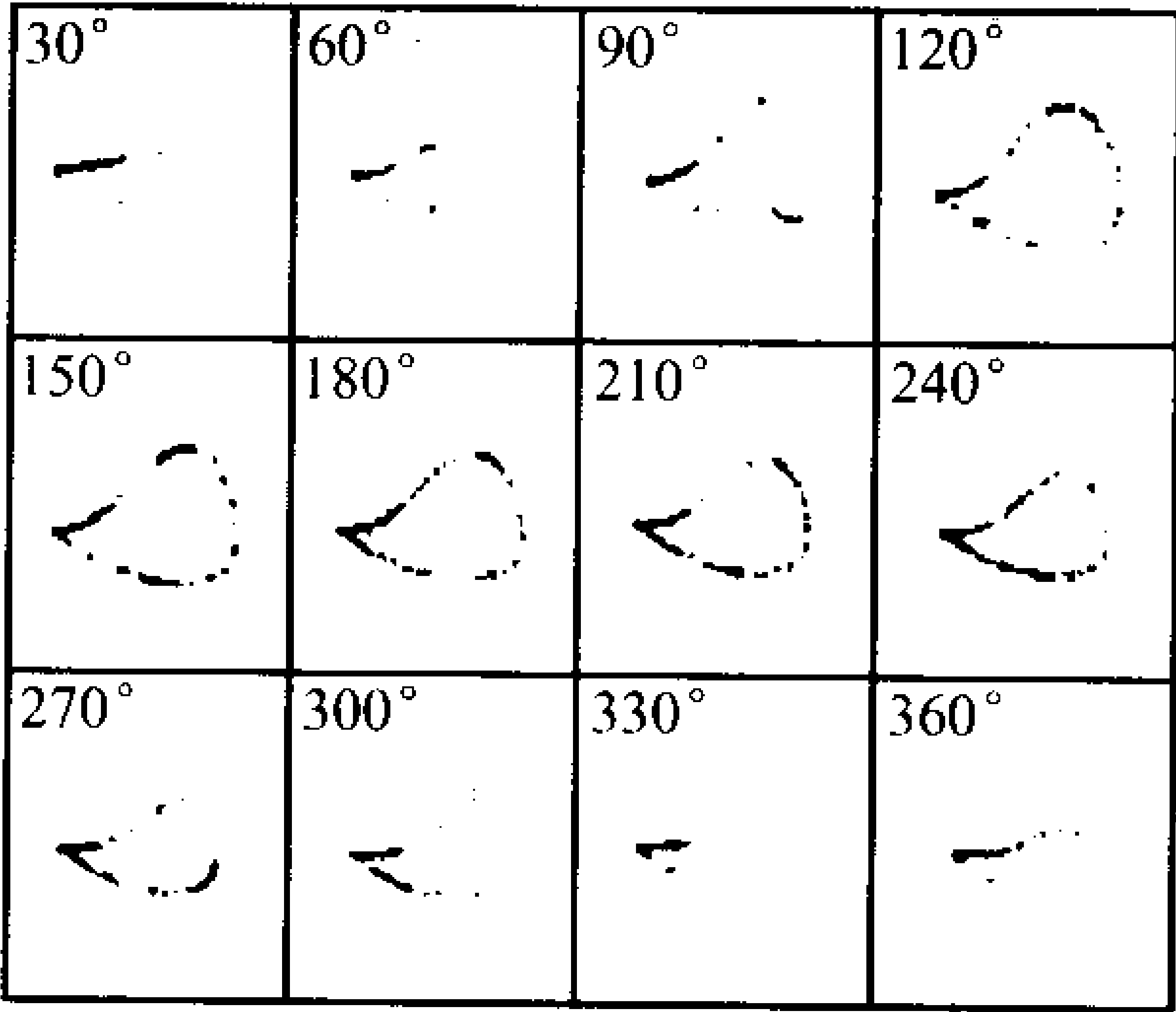
(b) 状态空间 $V\times\dfrac{\mathrm{d}V}{\mathrm{d}t}$ 内轨迹的横切面

图 3.2.7　乌贼巨轴突的间歇混乱的混沌响应

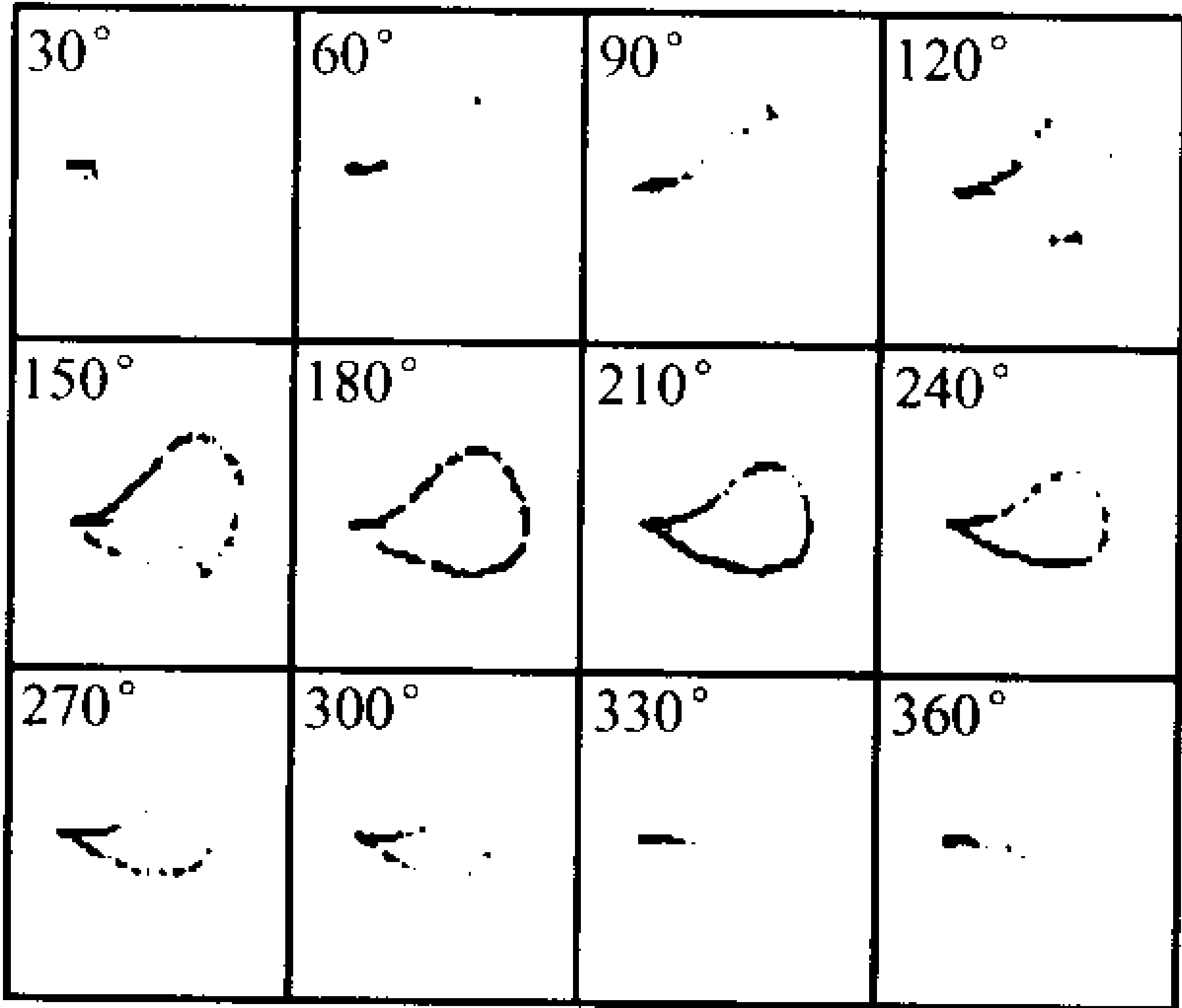
且也显示了从周期响应及准周期响应向混沌响应的过渡倍周期分叉道路与环面崩溃的道路，从中可见真实的生物神经膜电位的响应特性构成了一个带有明显混沌的不完整的魔鬼楼梯．而单一脉冲的电流刺激产生的活动电势并不服从在空间箝制条件下的全或无定律，对于空间均匀的情形，输入 —— 响应曲线并非不连续地全或无，而是连续缓变的，即响应特性不能用前述模型中的跳跃函数 f_{01} 来描述．虽然空间箝制条件看起来似乎过于人为，但值得注意的是活动电位实际上都是在一个成为轴突小点或者触发地带的



(a) 准周期响应



(b) 变形



(c) 混沌响应

图 3.2.8　乌贼巨轴突从准周期响应向混沌响应的过渡

轴突的局部开始的. 触发地带的长度, 用活动电位的自激振荡状态中乌贼巨轴突的实验估计, 约为 1mm, 与空间均匀性长度近似一致.

这个实验启发我们必须修正前面的 Caianiello 模型及 Nagumo–Sato 模型. 下面介绍的 Aihara 混沌神经网络模型就是一种很好的修正模型.

3.2.5 Aihara 神经元简化现实性模型

1990 年日本学者 Aihara 指出, "生物神经元具有通常的人工神经元所不具备的典型特性之一就是在单个神经元中可从实验上观察到混沌行为, 例如当给静息神经细胞膜加上周期性刺激时, 无论是对乌贼巨轴突进行实验, 还是对它的模型 Hodkin–Kuxley 方程求取数值解, 都可以观察到混沌行为. 但是如果每个神经元的模型都要建立在离子通道的基础上, 那么单个单元的模型就已相当复杂, 再要研究由它们构成的网络就更困难了. 下面将介绍由 Aihara (1990 年) 提出的一种神经元的简化现实性模型, 这种模型考虑生物神经元的主要外特征, 而不考虑其下一层次离子通道性质, 但它们仍能定性地再现生物神经元的混沌行为.

在前面的 Nagumo–Sato 模型中的跳跃函数 H 用连续递增、值域在 $[0,1]$ 中的 S 形函数 f 代替. 假定神经元在 $t+1$ 时刻的输出为 $x(t+1)$, 它取值在 $[0,1]$ 中以代替前面模型中状态变量仅取 $0, 1$ 两值的状态变量, 此时的数学模型为

$$x(t+1) = f\left(A(t) - \alpha\sum_{r=0}^{t} k^r x(t-r) - \theta\right), \tag{3.2.10}$$

比较更一般的模型形式为

$$x(t+1) = f\left(A(t) - \alpha\sum_{r=0}^{t} k^r g(x(t-r)) - \theta\right), \tag{3.2.11}$$

其中 g 函数是描述神经元输出与不应度大小之间的关系的函数.

定义 (3.2.10) 模型中的神经元的内部状态为

$$y(t+1) = A(t) - \alpha \sum_{r=0}^{t} k^r x(t-r) - \theta,$$

则 (3.2.10) 可以化简为

$$\begin{aligned} y(t+1) &= ky(t) - \alpha f(y(t)) + a(t), \\ x(t+1) &= f(y(t+1)), \end{aligned}$$

其中 $a(t)$ 仍为 (3.2.5) 所定义.

可以定义平均激发率

$$\rho = \lim_{n\to\infty} \sum_{t=0}^{n-1} \frac{l(x(t))}{n},$$

其中 l 为一函数, 它表示轴突的整形动力学, 在活动电位的传递过程中, 对轴突响应波形具有整形作用, 这时

$$l(x) = \begin{cases} 1, & x \geqslant 0.5, \\ 0, & x < 0.5. \end{cases}$$

要注意的是, 和空间箝制条件不同, 若轴突的长度充分长, 对活动电位沿轴突的传播, 全或无定律成立.

图 3.2.9 给出了模型 (3.2.10) 的响应特性图, 其中 $k = 0.5$, $\alpha = 1.0$, $f(y) = \dfrac{1}{1+\mathrm{e}^{-y/0.04}}$, $y(0) = 0.5$.

图 3.2.9(a) 给出了内部状态 y 依赖于分支参数 a 的分支图.

图 3.2.9(b) 给出了 Lyapunov 指数随分支参数 a 的变化图.

图 3.2.9(c) 给出了平均激发率随分支参数 a 的变化呈魔鬼楼梯状的图形. 这些响应图表明乌贼区轴突实验中观察到的交错的周期 —— 混沌响应序列被复制了.

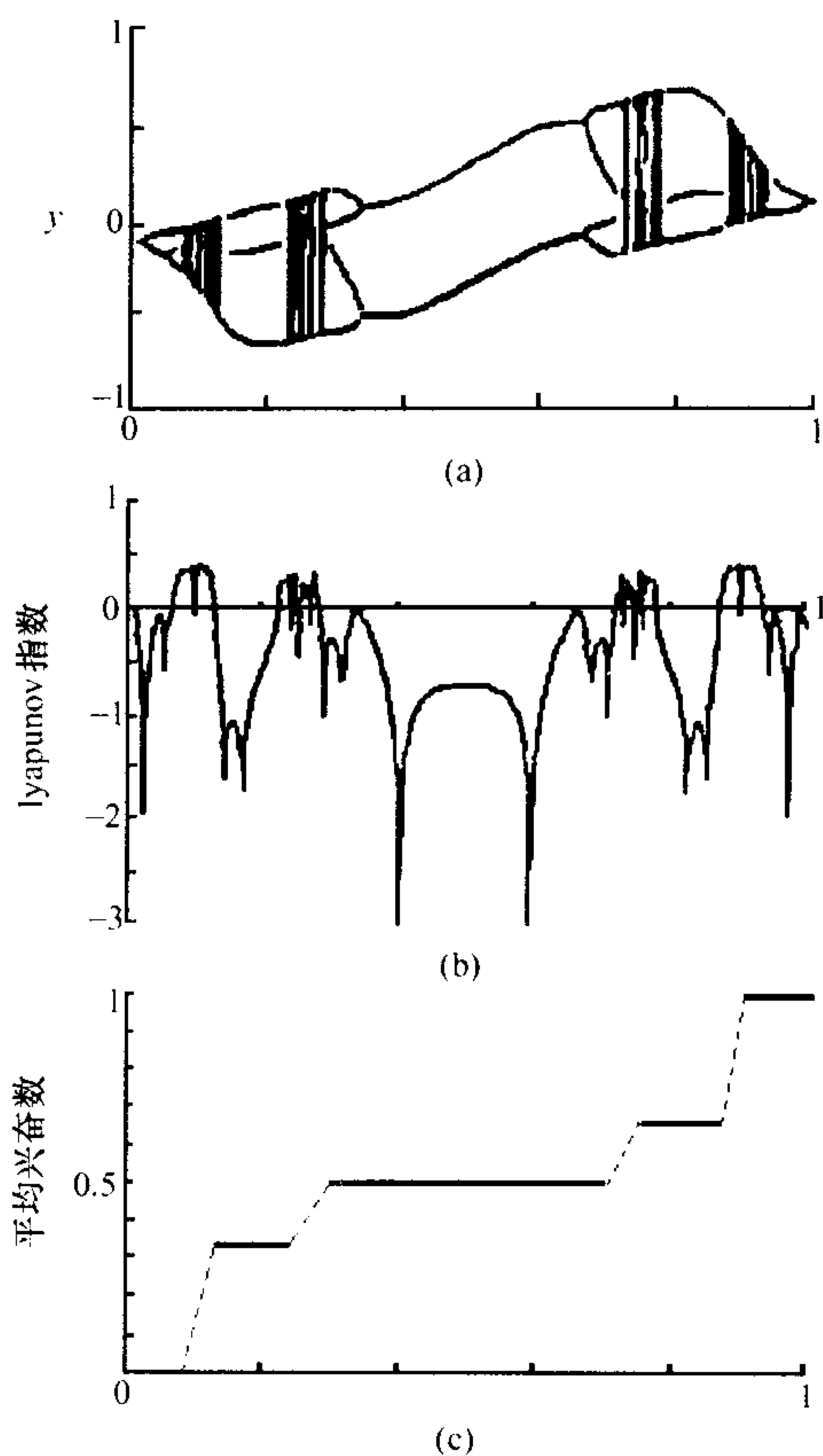


图 3.2.9　Aihara 神经元模型的分支响应图

(a) y 随 a 变化分支图; (b) Lyapunov 指数随 a 变化图; (c) 平均兴奋数随 a 变化图

如果在 f 中引进陡度参数 ε, 即定义

$$f(y)=\frac{1}{1+\mathrm{e}^{-y/\varepsilon}},$$

图 3.2.10 显示了 Aihara 神经元模型的响应示例. 图 3.2.10(a) 为 1

对 2 周期响应，其中 $k = 0.7$, $\alpha = 0.5$, $\varepsilon = 0.02$, 图 3.2.10(b) 为混沌响应，其中 $k = 0.7$, $\alpha = 0.35$, $\varepsilon = 0.02$.

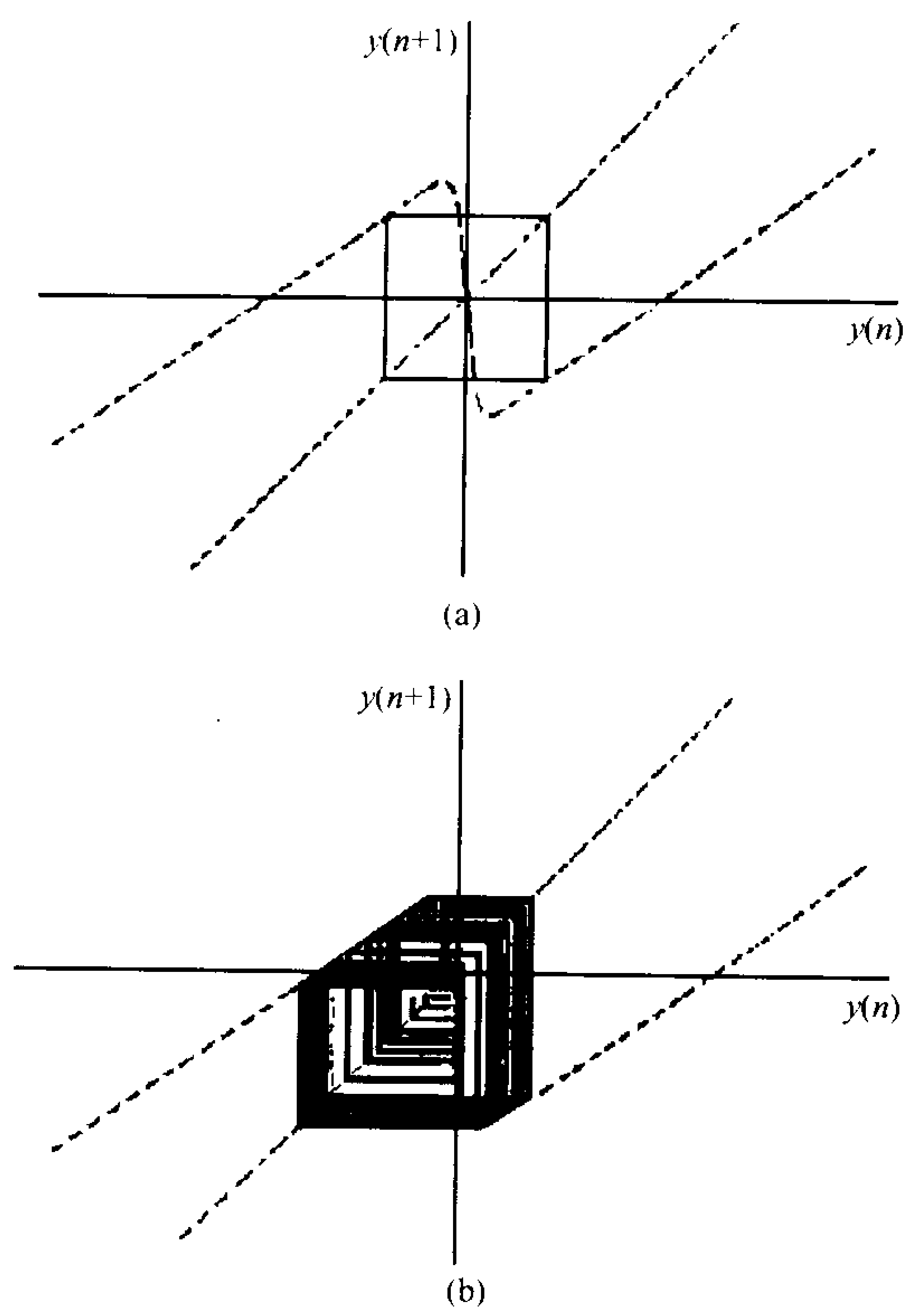


图 3.2.10　Aihara 神经元模型的响应示例

(a) 1 对 2 周期响应 ($k = 0.7$, $\alpha = 0.5$, $\varepsilon = 0.02$); (b) 混沌响应 ($k = 0.7$, $\alpha = 0.35$, $\varepsilon = 0.02$)

在参数空间 $a \times k$ 中方程 (3.2.10) 解的分类在图 3.2.11 中表示. 图中阴影区域对应于混沌解. 从图 3.2.11 可证实, 方程 (3.2.10) 和 (3.2.11) 的 Aihara 混沌神经元模型在广阔的参数空间区域内具有混沌解.

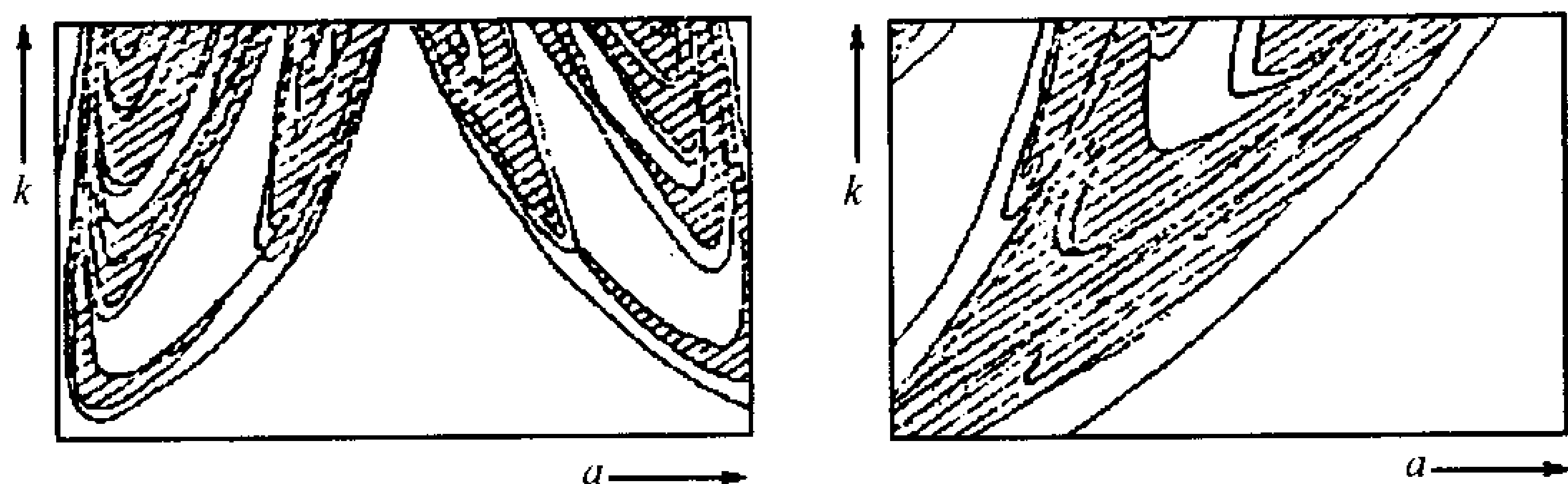


图 3.2.11　参数空间 $a \times k$ 内方程 (3.2.10) 解的分类

§3.3　离散的单层反馈型神经网络的动力学模型与问题

N 个神经元排列成一个单层，如果它构成全反馈的网络，这其中的每个神经元的输出都与其它神经元的输入相连，又整个网络的输入与输出的神经元数是相同的，都为 N. 记第 j 个神经元在 t 时刻的输出为 $x_j(t)$, 则在时刻 t 的 N 个神经元的输出为向量 $x(t) = (x_1(t), x_2(t), \cdots, x_N(t))^{\mathrm{T}} \in \mathrm{R}^N$, 而第 i 个神经元在时刻 t 的内部状态为 $y_i(t)$, 即 t 时刻 N 个神经元的内部状态为向量 $y(t) = (y_1(t), y_2(t), \cdots, y_N(t))^{\mathrm{T}} \in \mathrm{R}^N$, 如果

$$x_i(t+1) = f(y(t)), \tag{3.3.1}$$

其中 f 可以是最简单的二值函数 H, 也可以是 S 形单调连续函数，那么得到离散的单层反馈型神经网络. 下面将介绍离散的 Hopfield 神经网络及 Aihara 的混沌神经网络.

3.3.1　离散的 Hopfield 神经网络

1. 离散的 Hopfield 神经网络的动力学模型

假定在 (3.3.1) 中的 f 取二值符号函数 sgn 形式，又第 i 个神经元在时刻 t 的内部状态

$$y_i(t) = \sum_{j=1}^{N} w_{ij} x_j(t) + I_i, \tag{3.3.2}$$

也就是 $x_i(t+1)$ 有如下两种形式：若 $y_i(t) \geqslant 0$, 则 $x_i(t+1)=1$, 即第 i 个神经元在 $t+1$ 时刻兴奋；若 $y_i(t)<0$, 则 $x_i(t+1)=0$, 即第 i 个神经元在 $t+1$ 时刻抑制. 写成统一形式，网络的数学形式按网络的工作方式可以写成如下异步和同步两种方式:

(1) 异步方式：如果在某一时刻 $t+1$, N 个神经元中按照随机方式也可按照预先指定的顺序来选择某一个神经元 (例如第 j 个神经元), 它的输出为

$$x_j(t+1)=\operatorname{sgn}(y_j(t)), \tag{3.3.3}$$

但是其它 $N-1$ 个神经元的输出保持不变

$$x_i(t+1)=x_i(t), \quad i \neq j. \tag{3.3.4}$$

(2) 同步方式：在任何时刻 $t+1$, N 个神经元中的每一个神经元的输出均按同一方式生成

$$x_i(t+1)=\operatorname{sgn}(y_i(t)), \quad i=1,2,\cdots,N. \tag{3.3.5}$$

这两种方式中的第 i 个神经元的内部状态 $y_i(t)$ 为

$$y_i(t)=\sum_{j=1}^{N} w_{ij} x_j(t)+I_i, \quad i=1,2,\cdots,N. \tag{3.3.6}$$

2. 上述模型的收敛与振荡的动力学行为的分析

当输入向量 $I=(I_1,I_2,\cdots,I_N)^{\mathrm{T}}$ 作为一个初值时，网络通过反馈演化，从网络的输出端得到一个向量 $x(t)=(x_1(t),x_2(t),\cdots,x_N(t))^{\mathrm{T}}$, 如果它是收敛的，即指存在某时刻 t, 使得

$$x(t+1)=x(t),$$

数学上称它为 **稳定的平衡点**. 如果它是振荡的，例如是二周期的周期振荡，即存在时刻 t, 使得

$$x(t+2)=x(t), \quad x(t+3)=x(t+1), \quad \cdots,$$

或者从某时刻 t 开始，输出 $x(t), x(t+1)$ 周期性重复振荡，

$$x(t), x(t+1), x(t), x(t+1), \cdots.$$

下面证明两个基本定理，分别对于异步网络与同步网络各种情形给出的输出 $x(t)$ 收敛及振荡的结果，所用方法是借助广义能量函数或者物理中铁磁材料的 Hamilton 函数作为 Lasalle 不变性原理中的 Lyapunov 函数.

在铁磁体自旋中只有两个方向，即铁磁分子自旋 $P_i \in \{1, -1\}$ $(i = 1, 2, \cdots, n)$，这种材料中的 Hamilton 函数为

$$H = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} J_{ij}P_iP_j - \sum_{i=1}^{n} H_iP_i, \tag{3.3.7}$$

其中 J_{ij} 表示第 i 个分子与第 j 个分子之间的作用，$J_{ij} = J_{ji}$，即 J 矩阵是对称的. H_i 为外加的随机场，整个物质的相互作用结果是使 Hamilton 函数 H 达到最小. 由于在离散的反馈型 Hopfield 神经网络中，输出 $x_i(t)$ 也是二值的量，而第 i 个神经元与第 j 个神经元相互作用由权 w_{ij} 表示，不妨也设满足 $w_{ij} = w_{ji}$，即权矩阵是对称的. 当然这个假定是人为的，对于人工神经网络的模型，工程上的应用是合理的、可行的，但是对于生物意义下就显得太"过分"了，与实际人脑中神经网络的连接方式是不太符合的. 相应地定义离散反馈的 Hopfield 神经网络的广义能量函数 (或称为 Lyapunov 函数) 为如下形式:

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}x_ix_j - \sum_{i=1}^{N} I_ix_i, \tag{3.3.8}$$

显见，由于 x_i, x_j 只能取 $+1$ 或 -1 二值，w_{ij}, I_i $(i, j = 1, 2, \cdots, N)$ 均为有界的，

$$|E| \leqslant \frac{1}{2}\sum_{i,j=1}^{N} |w_{ij}| + \sum_{i=1}^{N} |I_i|.$$

从任意一个初始状态出发，如果每次迭代都能满足相邻两个时刻的能量函数的变化 $\Delta E \leqslant 0$，那么网络的能量随时间变化越来

越小, 最后趋向于极限集中的最大的不变集, 可能是平衡点集, 也可能是极限环周期解集, 在这些集合上的点满足 $\Delta E = 0$. 它的物理意义是, 在那些渐近稳定的不动点集及周期点集的吸引域内, 当状态离吸引子越远, 则能量越大, 当状态离吸引子越接近, 则能量越小. 而能量 E 的单调下降, 说明状态的运动从远离吸引子的地方跑到了吸引子上, 最后网络到达稳定.

下面分异步与同步两种方式讨论动力学行为:

(1) 异步方式

定理 3.3.1 如果网络的工作方式是异步的, 且满足权矩阵对称与对角元非负的条件, 则网络的终态是收敛的, 且趋向于渐近稳定的不动点.

证 对于离散反馈型的 Hopfield 神经网络的第 j 个神经元, 它的相邻输出时间的变化为

$$\Delta x_j = x_j(t+1) - x_j(t) = \begin{cases} 0, & 当\ x_j(t+1) = x_j(t)\ 时, \\ +2, & 当\ x_j(t+1) = 1,\ x_j(t) = -1\ 时, \\ -2, & 当\ x_j(t+1) = -1,\ x_j(t) = 1\ 时. \end{cases}$$

而相应的 Lyapunov 函数为

$$E = -\frac{1}{2}\sum_{i,j=1}^{n} w_{ij}x_i x_j - \sum_{i=1}^{n} I_i x_i,$$

相邻时间的 Lyapunov 函数的变化为

$$\begin{aligned} \Delta E = & -\frac{1}{2}\sum_{\substack{i,j=1\\ j\neq i}}^{n} w_{ij}x_i\Delta x_j - \frac{1}{2}\sum_{\substack{i,j=1\\ j\neq i}}^{n} w_{ji}x_i\Delta x_j \\ & -w_{jj}[x_j^2(t+1) - x_j^2(t)] - I_j\Delta x_j \\ = & -\left(\sum_{\substack{i,j=1\\ j\neq i}}^{n} w_{ij}x_i + I_j\right)\Delta x_j \end{aligned}$$

$$= -\left(\sum_{i=1}^{n} w_{ij}x_i + I_j\right)\Delta x_j + w_{jj}x_j\Delta x_j,$$

这里用到了 $w_{ij} = w_{ji}$ 以及 $x_j^2(t+1) - x_j^2(t) = 1 - 1 = 0$.

如果 $w_{jj} = 0$, 那么 $\Delta E = -y_j(t)\Delta x_j \leqslant 0$. 对于满足 $\Delta E = 0$ 的集合中的点有 2 种可能性：或 $\Delta x_j = 0$, 或 $y_j(t) = 0$, 加上最大不变集的条件得 $\Delta E = 0$ 的集合中的最大不变集为 $\Delta x_j = 0$, 即不动点 $x_j(t+1) = x_j(t)$, 且该不动点是稳定吸引子.

如果 $w_{jj} > 0$, 那么 $x_j(t)[x_j(t+1) - x_j(t)]$ 可能取的值为 0 或 -2, 所以 $\Delta E \leqslant -y_j(t)\Delta x_j \leqslant 0$. 又 $\Delta E = 0$ 中最大不变集为 $\Delta x_j = 0$, 即不动点 $x_j(t+1) = x_j(t)$ 为稳定吸引子. 证毕.

(2) 同步方式

定理 3.3.2 如果网络的权矩阵 W 是对称的, 则网络的终态动力学行为可能是收敛的, 即趋向于渐近稳定的不动点, 也可能是稳定的周期 2 解振荡的.

证 作 Lyapunov 函数

$$E \triangleq -\frac{1}{2}\sum_i\sum_j w_{ij}x_i(t+1)x_j(t) - \frac{1}{2}\sum_i I_i[x_i(t) + x_i(t+1)]$$
$$= -\frac{1}{2}x^{\mathrm{T}}(t+1)Wx^{\mathrm{T}}(t) - \frac{1}{2}I^{\mathrm{T}}[x(t+1) + x(t)], \tag{3.3.9}$$

其中 W 是矩阵 (w_{ij}) $(i,j = 1,2,\cdots,N)$, $I = (I_1, I_2, \cdots, I_N)^{\mathrm{T}}$, $x(t) = (x_1(t), x_2(t), \cdots, x_n(t))^{\mathrm{T}}$. 易知

$$\begin{aligned}\Delta E &= -\frac{1}{2}x^{\mathrm{T}}(t+1)Wx(t) - \frac{1}{2}I^{\mathrm{T}}[x(t+1) + x(t)]\\ &\quad - \left\{-\frac{1}{2}x^{\mathrm{T}}(t)Wx(t-1) - \frac{1}{2}I^{\mathrm{T}}[x(t) + x(t-1)]\right\}\\ &= -\frac{1}{2}[x^{\mathrm{T}}(t)W][x(t+1) - x(t-1)] - \frac{1}{2}I^{\mathrm{T}}[x(t+1) - x(t-1)]\\ &= -\frac{1}{2}[x^{\mathrm{T}}(t)W + I^{\mathrm{T}}][x(t+1) - x(t-1)]\\ &= -\frac{1}{2}y^{\mathrm{T}}(t)[x(t+1) - x(t-1)].\end{aligned}$$

注意到

$$y_i(t)x_i(t+1) = y_i(t)\cdot \mathrm{sgn}[y_i(t)] \geqslant 0,$$

当 $y_i(t) < 0$, 即 $x_i(t+1) = -1$, 故 $x_i(t+1) - x_i(t-1) = 0$ 或 -2; 当 $y_i(t) > 0$, 即 $x_i(t+1) = 1$, 故 $x_i(t+1) - x_i(t-1) = 0$ 或 2. 从而有 $y_i(t)[x_i(t+1) - x_i(t-1)] \geqslant 0$, 即 $\Delta E \leqslant 0$. 考虑满足 $\Delta E = 0$ 的集合中的点有三种可能性: 或 $x(t+1) = x(t-1)$; 或 $y(t) = 0$; 或 $x(t+1) \neq x(t-1)$, $y(t) \neq 0$, $y^{\mathrm{T}}(t)[x(t+1) - x(t-1)] = 0$. 仅有 $x(t+1) = x(t-1)$ 满足的点是不变集的点, 故由 Lasalle 不变性原理知道, 解的轨道趋向于满足 $x(t+1) = x(t-1)$ 的点集, 它或者是收敛的不动点 $x(t+1) = x(t) = x(t-1)$, 或者是周期 2 解 $x(t+1) = x(t-1)$. 前者是不动点吸引子, 即终态动力学行为是收敛的, 后者是周期 2 的极限环吸引子, 即终态动力学行为是周期 2 振荡的. 证毕.

3. 上述的收敛与振荡的动力学行为与联想记忆的关系

对于 m 个不同的输入样本向量 $V^1, V^2, \cdots, V^m$, 联想记忆问题就是一方面希望 "记住" 所存储的样本集, 即当每一个输入的样本为系统的初值, 经过不断演化最后的终态又变为自己, 即每一个样本向量都是网络最终的收敛的稳定不动点; 另一方面还要求每个样本周围都有适当大的 "吸引域", 这样才能对诸如模糊、畸变、不完整的输入有联想识别的能力. 虽然网络能够在样本集输入时收敛到自己, 但并不说明整个系统一定是稳定的, 即某些非样本矢量输入时, 系统可能不是收敛到稳定的不动点 —— 需要的记忆样本集, 而是呈现振荡或混沌奇异吸引子比较复杂的动力学行为. 前面一段对同步方式网络的分析就说明了这种可能性.

对上述联想记忆的两个方面的要求在一些适当的假设条件下, 即对所需记忆的样本向量给出一定的限制下, 通过选择适当的权矩阵可以达到目的. 以下讨论假设 $I = 0$, 模型取为同步方式.

如果仅仅要求 "记住" 所存的样本集, 那么只需要求权矩阵满足

$$WV^k = \beta V^k \quad (k = 1, 2, \cdots, m),$$

其中

$$V^k = (V_1^k \quad V_2^k \quad \cdots \quad V_n^k) \in \mathrm{R}^n,$$

$$W = \begin{pmatrix} w_{11} & w_{21} & \cdots & w_{n1} \\ \vdots & \vdots & & \vdots \\ w_{1n} & w_{2n} & \cdots & w_{nn} \end{pmatrix},$$

$$\beta = \begin{pmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_n \end{pmatrix}, \quad \beta_i > 0 \quad (i = 1, 2, \cdots, n).$$

将 m 个 V^k 联合起来写成矩阵方程组形式

$$W(V^1 \quad V^2 \quad \cdots \quad V^m) = \beta(V^1 \quad V^2 \quad \cdots \quad V^m).$$

记

$$Z = (V^1 \quad V^2 \quad \cdots \quad V^m) \in \mathrm{R}^{n\times m},$$
$$B = (\beta_1 V^1 \quad \beta_2 V^2 \quad \cdots \quad \beta_m V^m) \in \mathrm{R}^{n\times m},$$

有 $WZ = B$. 易知 $Z^{\mathrm{T}}W^{\mathrm{T}} = B^{\mathrm{T}}$, 记 W^{T} 中的第 j 列为 W_j, B^{T} 中的第 j 列为 B_j, 可得到下面具有 n 个未知数 $w_{1j}, w_{2j}, \cdots, w_{nj}$ 的 m 个方程

$$Z^{\mathrm{T}}W_j = B_j \quad (j = 1, 2, \cdots, n).$$

如果 $m \leqslant n$, Z 的秩为 m, 即样本集 m 个向量为线性无关的, 那么总可以由这 m 个方程解出解 $W_j = (w_{1j}, w_{2j}, \cdots, w_{nj})^{\mathrm{T}}$ $(j = 1, 2, \cdots, n)$, 使得当输入 $V^1, V^2, \cdots, V^m$ 后, 网络最终收敛在这些样本集.

对于比较高的联想记忆的要求, 已经有很多研究工作, 典型的方法有外积型权的设计方法. 下面将介绍一种建立在 Lyapunov 函数的特征分析基础上的方法, 它有较多的优点.

(1) Lyapunov 函数的特征分析

考虑一类离散的同步 Hopfield 神经网络 ($I=0, W=W^{\mathrm{T}}$), 它的能量函数即 Lyapunov 函数取为

$$E=-\frac{1}{2}x^{\mathrm{T}}Wx. \tag{3.3.10}$$

由 W 对称性可得到 W 的如下形式的表达式:

$$W=\sum_{i=1}^{M}\lambda_i q_i q_i^{\mathrm{T}}, \tag{3.3.11}$$

其中 λ_i 是矩阵 W 的特征值, 而 q_i 是 $N\times 1$ 的特征向量. 所有的特征向量均是相互正交的, 而且这些特征向量的 Euclid 范数为 1, 即

$$q_j^{\mathrm{T}}q_i=\begin{cases}1, & i=j,\\ 0, & \text{其它}.\end{cases} \tag{3.3.12}$$

由于权矩阵 W 是对称的, 所以所有的特征值均为非负. 当然, 有的特征值为退化的, 即一个特征值对应于几个特征向量, 在这种情况下, 每个特征值对应于一个子空间. 如果矩阵的特征值为退化的, 则相应的特征向量构成的子空间为核空间, 因此具有核空间是 Hopfield 网络的内在特点是 $M<N$.

方程 (3.3.11) 是矩阵 W 的特征分解, 相应的状态向量 x 的特征表示为

$$x=\sum_{i=1}^{M}\gamma_i q_i+x_{\mathrm{null}}, \tag{3.3.13}$$

其中 $\gamma_i q_i$ 是向量 x 中的沿着特征向量 q_i 方向的一部分, 而 x_{null} 是向量 x 中位于核空间中的一部分. 于是将 (3.3.11), (3.3.13) 代入 (3.3.10) 得到以下 Lyapunov 函数

$$E=-\frac{1}{2}\sum_{i=1}^{M}\lambda_i\gamma_i^2, \tag{3.3.14}$$

其中应用了方程 (3.3.12) 对非零特征值的标准正交的关系，对核空间有

$$x_{\text{null}}^{\mathrm{T}} q_i = 0 \quad (i = 1, 2, \cdots, M). \tag{3.3.15}$$

对 (3.3.14) 式的展开观察后，现在可以认为我们的方法是使能量函数 E 最小化. 我们可以按如下方式移动在相平面中的状态向量 x. 如果特征值 λ_i 是负数，则使得 γ_i 为 0, 否则 λ_i 为正数，增大 γ_i 的值. 注意 $\gamma_i q_i$ 是向量 x 中沿着特征向量 q_i 方向的一部分，相应于特征值 λ_i, 而 q_i 和 λ_i 均与突触权矩阵 W 有关.

(2) 学习规则

在联想记忆中，当输入有偏差的模式后，仍能使之给出已经存储的记忆模式. 记存储的记忆模式个数为 M, 每个模式由值为 ± 1 的 N 个元素构成. 相应的，一个离散的 Hopfield 模型被设计成具有 N 个神经元，而权矩阵的确定可以由要进行存储的学习方法决定，这种方法有两种特点:

1) 记忆向量必须是网络的稳定的不动点;

2) 当初值落在某一个不动点的吸引域中，这一状态必定被吸引到一个恰当的已经存储的记忆向量.

记 $\overline{x}$ 是记忆向量，根据第一点，其必须是不动点. 于是由不动点的定义，要求

$$\overline{x} = \operatorname{sgn}(W\overline{x}). \tag{3.3.16}$$

根据 (3.3.13), 可以得到

$$\overline{x} = \sum_{i=1}^{M} \overline{\gamma}_i q_i + \overline{x}_{\text{null}}, \tag{3.3.17}$$

其中 $\overline{\gamma}_i q_i$, $\overline{x}$, q_i, $\overline{x}_{\text{null}}$ 的定义与之间的关系与前面的类似. 于是结合 (3.3.11) 和 (3.3.17), 并根据 (3.3.12) 和 (3.3.15) 有如下结果

$$W\overline{x} = \sum_{i=1}^{M} \lambda_i \overline{\gamma}_i q_i. \tag{3.3.18}$$

将 (3.3.17), (3.3.18) 代入 (3.3.16) 后，有

$$\sum_{i=1}^{M} \overline{\gamma}_i q_i + \overline{x}_{\text{null}} = \operatorname{sgn}\left(\sum_{i=1}^{M} \lambda_i \overline{\gamma}_i q_i\right). \tag{3.3.19}$$

(3.3.19) 保证对任何向量均成立，则必须加以下条件:

$$\overline{x}_{\text{null}} = 0 \tag{3.3.20}$$

与

$$\lambda_i = \lambda (i = 1, 2, \cdots, M) \quad \text{且} \quad \lambda > 0. \tag{3.3.21}$$

换言之，我们可以得到如下结论 (Aiyer 等， 1990):

1) 当 $\overline{x}_{\text{null}}$ 等于零向量时，则突触权矩阵 W 的核空间与所有的记忆向量正交.

2) 如果 $\overline{x}_{\text{null}}$ 为 0, 那么从 (3.3.17) 式可知所有的记忆向量可以由和式完全地给出，即 $\overline{x} = \sum_{i=1}^{M} \overline{\gamma}_i q_i$. 也就是说，矩阵 W 的特征向量可张成由记忆向量生成的子空间，这只是将第一点再重申一遍.

3) 对于 $\lambda_i = \lambda\ (i = 1, 2, \cdots, M)$, 矩阵 W 必有一个正的退化的特征值 $\lambda > 0$, 其相应的特征向量空间是记忆向量张成的子空间.

那么最简单的学习方法就是外积规则 (outer product rule). 记 $\xi_1, \xi_2, \cdots, \xi_M$ 表示 M 个基本的学习向量，每个向量由 N 个元素构成，于是权矩阵可以由外积规则给出

$$W_M = \sum_{i=1}^{M} \xi_i \xi_i^{\mathrm{T}}. \tag{3.3.22}$$

对 $N \times 1$ 的学习向量作如下假设:

1) 学习向量的元素取值为 ±1, 则 ξ_i 的内积为

$$\xi_i^{\mathrm{T}} \xi_i = N \quad (i = 1, 2, \cdots, M); \tag{3.3.23}$$

2) 记忆向量是线性无关的.

如果记 $\mathcal{M}$ 表示由学习向量 $\xi_1, \xi_2, \cdots, \xi_M$ 张成的子空间, $\mathcal{N}$ 是与 $\mathcal{M}$ 正交的子空间, 于是 $\mathcal{N}$ 是矩阵 W_M 的核空间. 记 $\overline{\xi}$ 与 $\overline{x}_{\text{null}}$ 表示分别位于子空间 $\mathcal{M}$ 与 $\mathcal{N}$ 中不动点 $\overline{x}$ 的两部分, 则有

$$\overline{x} = \overline{\xi} + \overline{x}_{\text{null}}. \tag{3.3.24}$$

此外由于 $\{\xi_i\}_{i=1}^M$ 是线性无关的, 那么 $\overline{\xi}$ 可以记成

$$\overline{\xi} = \sum_{\alpha=1}^{M} \overline{\gamma}_\alpha \xi_\alpha. \tag{3.3.25}$$

于是, 利用 (3.3.22), (3.3.24), (3.3.25), 可以表示出矩阵的乘积 $W_M\overline{x}$ 的形式如下:

$$W_M\overline{x} = \left(\sum_{\alpha=1}^{M} \xi_\alpha \xi_\alpha^{\mathrm{T}}\right) \left(\sum_{\beta=1}^{M} \overline{\gamma}_\beta \xi_\beta + \overline{x}_{\text{null}}\right). \tag{3.3.26}$$

下面给出如下 4 点注释:

1) 子空间 $\mathcal{M}$ 和 $\mathcal{N}$ 是相互垂直的, 即对所有的 α 成立

$$\xi_\alpha^{\mathrm{T}} \overline{x}_{\text{null}} = 0; \tag{3.3.27}$$

2) 由 (3.3.23) 式知内积 $\xi_\alpha^{\mathrm{T}} \xi_\alpha = N$;

3) 在 (3.3.25) 式中, 将 α 改记为 β, 则有

$$\overline{\xi} = \sum_{\beta=1}^{M} \overline{\gamma}_\beta \xi_\beta;$$

4) 最后, 引进噪声向量

$$\xi_{\text{noise}} = \sum_{\substack{\alpha=1 \\ \alpha \neq \beta}}^{M} \sum_{\beta=1}^{M} \overline{\gamma}_\beta \xi_\alpha \xi_\alpha^{\mathrm{T}} \xi_\beta. \tag{3.3.28}$$

将 (3.3.26) 式右端打开, 并注意到以上 4 点注释后, 有

$$W_M\overline{x} = N\overline{\xi} + \xi_{\text{noise}}. \tag{3.3.29}$$

这个方程清晰地说明了对于权矩阵 W_M 有 1 个等于 N 的退化的特征值，而噪声项 ξ_{noise} 必须为 0. 若噪声 $\xi_{\text{noise}} = 0$, 必须满足所有的学习向量互相正交，即

$$\xi_\alpha^{\mathrm{T}} \xi_\beta = 0 \quad (\alpha \neq \beta), \tag{3.3.30}$$

也就是 (3.3.22) 式的外积规则仅仅在满足学习向量相互正交时，才可以使用.

很多人所作的研究表明，当神经元个数 N 大于学习向量的个数 M, 且学习向量是随机地选自于一个较大规模的总体时，则学习向量相互正交的概率非常高，所以在概率论意义下，(3.3.20) 与 (3.3.21) 是可以满足的.

联想记忆

作为具有联想记忆功能的 Hopfield 网络，必须具备将与学习向量在海明距离下非常接近的试探向量 ξ_{probe} 恢复成相应的学习向量的能力，其中海明距离是指两个向量不同的存储单元的个数. 特别地，试探向量是由学习向量作了适当的偏差而得到的，于是 Hopfield 网络也可以被认为是一种纠错的网络.

先前，我们应用特征分析说明了沿着矩阵 W 负特征值的特征向量的状态向量的所有的元素均为 0, 同时指出了，若矩阵 W_M 是由外积规则给出的，则子空间 $\mathcal{N}$ 必然是 W_M 的核空间，于是其对应于矩阵 W_M 的零特征值. 这似乎表明，为了在子空间 $\mathcal{N}$ 中插入 1 个负特征值，可以对矩阵 W_M 作修正，于是对网络而言，完全可以提供一种有限的纠偏方法，使得除去位于这一空间中向量 ξ_{probe} 的一部分. 也就是说，这种纠偏技术的运用可以通过对位于子空间 $\mathcal{N}$ 中的元素的加法纠正由于向量 ξ_{probe} 的偏差而产生的错误.

将一个负特征值插入子空间 $\mathcal{N}$, 最简单的方法是从矩阵 W_M 中减去单位阵 I 与数的乘积，即

$$W = W_M - MI = \sum_{i=1}^{M} \xi_i \xi_i^{\mathrm{T}} - MI. \tag{3.3.31}$$

这一修正使得矩阵 W 对角线上的元素均变成了 0, 而这一修正的物理意义是 Hopfield 网络没有了自身到自身的闭环. 记向量

$$x = \bar{\xi} + x_{\text{null}}, \tag{3.3.32}$$

其中 $\bar{\xi}$ 是 x 在子空间 $\mathcal{M}$ 中的射影, 即

$$\bar{\xi} = \sum_{\alpha=1}^{M} \bar{\gamma}_\alpha \xi_\alpha, \tag{3.3.33}$$

而 x_{null} 是 x 在子空间 $\mathcal{N}$ 中的射影. 然后利用 (3.3.33) 式及导出 (3.3.29) 式的方法, 给出矩阵 W 的 Lyapunov 函数

$$E = -\frac{1}{2}x^{\mathrm{T}}Wx = -\frac{1}{2}(N-M)\|\bar{\xi}\|^2 - \frac{1}{2}\bar{\xi}\xi_{\text{noise}} + \frac{1}{2}M\|x_{\text{null}}\|^2, \tag{3.3.34}$$

其中 $\|\bar{\xi}\|$ 和 $\|x_{\text{null}}\|$ 表示 Euclid 意义下的范数. 纠错能力至少要求学习向量可以被修正为其自身, 所以学习向量必然是不动点. 正如先前所指出的, 只要从正交向量集合中选择学习向量, 所有条件均能成立, 而这种选择必然使 $\xi_{\text{noise}} = 0$, 所以 (3.3.34) 式可以简化为

$$E = -\frac{1}{2}(N-M)\|\bar{\xi}\|^2 + \frac{1}{2}M\|x_{\text{null}}\|^2. \tag{3.3.35}$$

这个式子表明能量函数在 $\|\bar{\xi}\| \to \infty$ 和 $\|x_{\text{null}}\| = 0$ 时达到最小. 在现实中, 由于能量函数在单位立方体的顶点处有最小值, 而极限 $\|\bar{\xi}\| \to \infty$ 与每个神经元的输出相抵制, 所以网络在立方体的一个顶点上达到稳定状态. 于是离散神经网络可以被认为是一种向量投影器, 将试探向量 ξ_{probe} 在子空间 $\mathcal{M}$ 上投影, 然后将得到的投影向量一步一步地移向最近的立方体的顶点上.

伪稳定状态

如果考虑 Hopfield 网络作为向量投影的信号释译, 那么就必须解释何谓伪稳定状态. 首先, 需要指出张成子空间 $\mathcal{M}$ 的基本学习向量构成了一个不动点集合, 这一集合相当于单位超立方体的

一部分顶点，而位于或离子空间 $\mathcal{M}$ 很近的另一些顶点是潜在的伪稳定状态. 更详细地，以下条件必须满足：

1) 如果超立方体的一个顶点位于子空间 $\mathcal{M}$ 中，而顶点在核空间 $\mathcal{N}$ 中的投影的范数远远小于在子空间 $\mathcal{M}$ 中投影的范数，即 $\|\overline{x}_{\text{null}}\| \ll \|\overline{\xi}\|$, 那么 $\text{sgn}(W\overline{x}) = \text{sgn}(\overline{\xi})$, 则顶点作为伪稳定状态的稳定性可以保证;

2) 如果顶点不在 $\mathcal{M}$ 中，但与之非常接近，即其在 $\mathcal{M}$ 上的投影 $\overline{\xi}$ 的元素并没有改变，那么 $\text{sgn}(\overline{x}) = \text{sgn}(\overline{\xi})$, 则伪稳定的顶点的稳定性可以保证.

条件 1) 比条件 2) 严格. 特别地，我们找到满足条件 1) 的点数非常少，而满足条件 2) 的点数关于学习向量的个数 M 呈指数递增. 此外，随着记忆向量的维数 N 关于 M 增加时，相应的伪稳定点的个数减小.

所以我们得出的结论是：在两种冲突的要求之间存在一种固有的权衡：(i) 保持学习向量在相空间中的稳定性；(ii) 减少伪稳定点. 前者要求网络的权矩阵 W 具有单个相应于子空间 $\mathcal{M}$ 的正退化特征值，这一要求反过来引起所有在 $\mathcal{M}$ 中的或距离 $\mathcal{M}$ 很近的顶点成为伪稳定点，而只有当伪稳定点的个数非常少以至于其与真正的向量投影器无法区分时， Hopfield 网络才能够成为一个联想记忆器.

3.3.2 Aihara 离散混沌神经网络动力学模型与问题

1. Aihara 离散混沌神经网络动力学模型

在离散的反馈 Hopfield 神经网络动力学模型的基础上，进一步考虑下面几个因素：

(1) 来自神经元本身的不应性影响 $g(x_i(t))$, 以及神经网络内部各神经元的反馈项 $h(x_j(t))$, h 为作用函数，其中也包含不应性影响并且通过权 w_{ij} 作用于这些反馈项;

(2) 来自神经元外部的输入项 $A_j(t)$ 也存在延迟一段时刻的作用，并通过类似权的形式施加对于 $x_i(t)$ 的影响;

(3) 在 (3.3.1) 中的 f 不取 H 跳跃函数而取为一般的 S 型单

调连续有界函数. 第 i 个神经元在时刻 t 的输出, 不是取 0, 1 两值而取 $[0,1]$ 或 $[-1,1]$ 的有限区间内的所有值.

$$x_i(t+1) = f\left(\sum_{j=1}^{M} v_{ij} \sum_{r=0}^{t} k^r A_j(t-r) + \sum_{j=1}^{N} w_{ij} \sum_{r=0}^{t} k^r h(x_j(t-r)) - \alpha \sum_{r=0}^{t} k^r g(x_j(t-r)) - \theta_i\right) \quad (i=1,2,\cdots,N), \tag{3.3.36}$$

其中 N 为网络中神经元个数, M 为外部输入数. 同样定义 f 函数作用的项为神经元的内部状态 $y_i(t+1)$, 即

$$y_i(t+1) = k y_i(t) + \sum_{j=1}^{M} v_{ij} A_j(t) + \sum_{j=1}^{N} w_{ij} h(f(y_j(t))) - \alpha g(f(y_j(t))) - \theta_i(1-k) \quad (i=1,2,\cdots,N), \tag{3.3.37}$$

$$x_i(t+1) = f(y_i(t+1)) \quad (i=1,2,\cdots,N). \tag{3.3.38}$$

为讨论简便起见, 设函数 g 和 h 均为恒等函数, 函数 f 取为 Logistic 函数

$$f(y) = \frac{1}{1+\mathrm{e}^{-y/\varepsilon}}. \tag{3.3.39}$$

外部输入不随时间变化, $A_j(t) = A_j$, 且定义常数为

$$a_i = \sum_{j=1}^{M} v_{ij} A_j - \theta_i(1-k). \tag{3.3.40}$$

这样 Aihara 神经网络动力学模型可以写成如下形式:

$$y_i(t+1) = k y_i(t) + \sum_{i=1}^{N} w_{ij} x_j(t) - \alpha x_i(t) + a_i \quad (i=1,2,\cdots,N), \tag{3.3.41}$$

$$x_i(t+1) = \frac{1}{1+\exp\left(-\dfrac{y_i(t+1)}{\varepsilon}\right)}, \tag{3.3.42}$$

称之为 Aihara 简化的混沌神经元网络模型.

2. Aihara 混沌神经网络的复杂动力学

因为在 §3.2 中已经介绍了如同 (3.3.12), (3.3.13) 神经网络中的每个神经元本身都可以产生混沌, 所以很自然, 有几个这种神经元组成的网络必然会表现为混沌动力学行为. 图 3.3.1 给出了神经元组成的 Aihara 神经元网络模型, 在每一个神经元的不应性系数 $\alpha = 1$ 假定下, 在 3 组不同的参数取值时得到的 $y_1(t)$ 与 $y_2(t)$ 相平面中的奇异吸引子图.

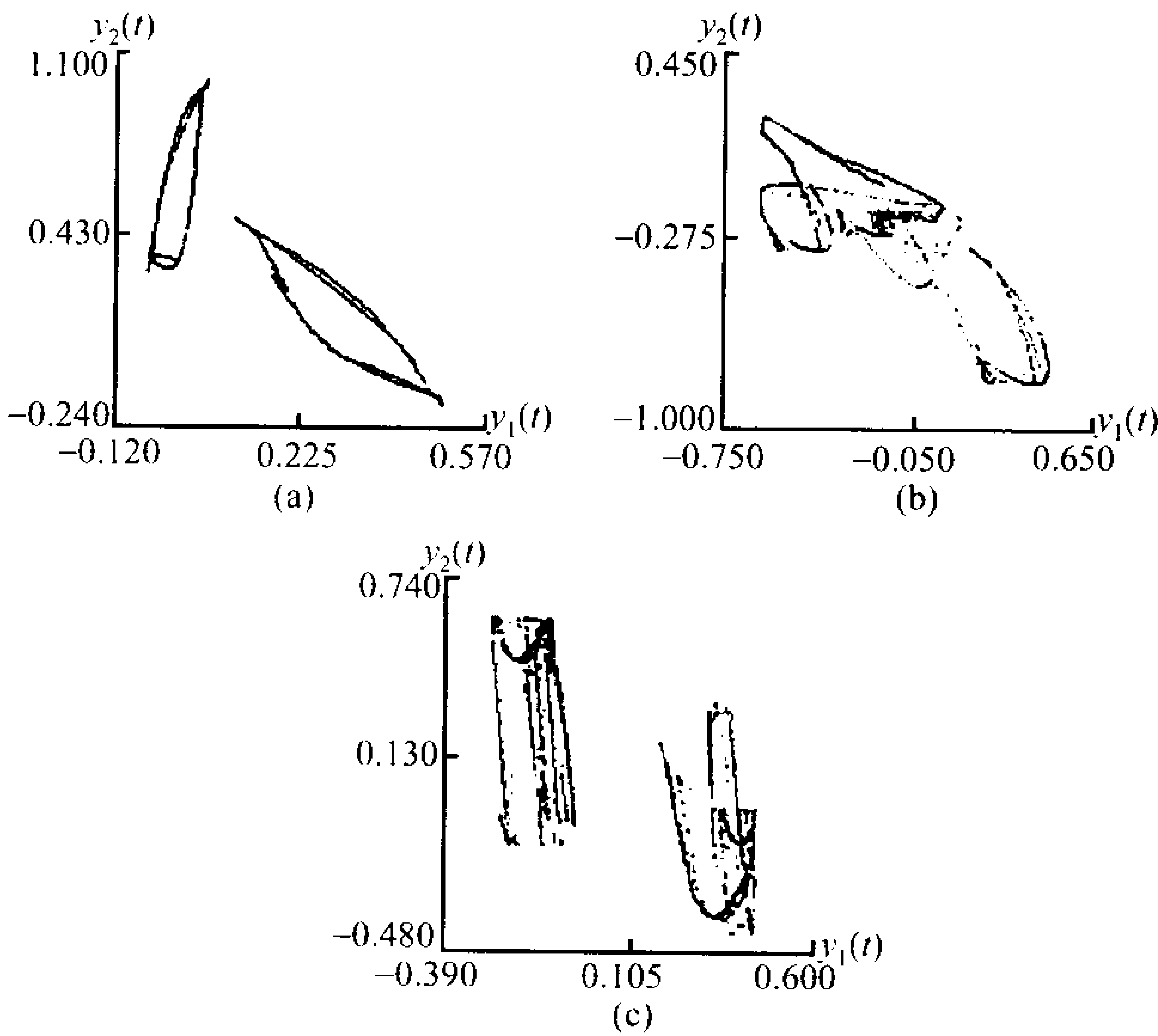


图 3.3.1 两个神经元组成的 Aihara 神经网络的混沌动力学

(a) $k = 0.5$, $\varepsilon = 0.06$, $a_1 = 0.8$, $a_2 = 0.5$, $w_{12} = 0$, $w_{21} = 0.7$; (b) $k = 0.9$, $\varepsilon = 0.04$, $a_1 = 0.2$, $a_2 = 0.04$, $w_{12} = 0.6$, $w_{21} = 0.4$; (c) $k = 0.8$, $\varepsilon = 0.03$, $a_1 = a_2 = 0.5$, $w_{12} = 0.1$, $w_{21} = 0.3$

3. 混沌奇异吸引子和动态记忆

下面介绍含 100 个混沌神经元的网络的计算机模拟及分析结果.

假定分支参数 $a_i = 0$, 连接权值 w_{ij} 通过 Hebb 学习规则建立

(若两个神经元的输出值为 x_i, x_j, 则它们之间连接权的变化 Δw_{ij} 与 x_i 及 x_j 成正比, 比例常数为 σ). 对于每个神经元, 均考虑其反馈项 (不含自反馈)、不应项和外部输入项的影响. 每个神经元采用相同的阈值函数, 即 Logistic 函数, 神经元输出是多值的 (本例中取 82 值), 并用 9×9 像素点组成的小方格表示. 方格内点数越多, 表示神经元输出值越大. 网络输出状态则用 10×10 个方格阵列表示. 在模拟过程中可以用计算机实时显示网络状态.

假定有 4 个标准模式 (每个模式含 10×10 个像素点) 作为记忆模式 (如图 3.3.2 的第一行中所显示的). 对网络进行训练, 即求得一个 100×100 的权值矩阵 w_{ij} $(i,j=1,2,\cdots,100)$, 其中 $w_{ii}=0$ $(i=1,2,\cdots,100)$.

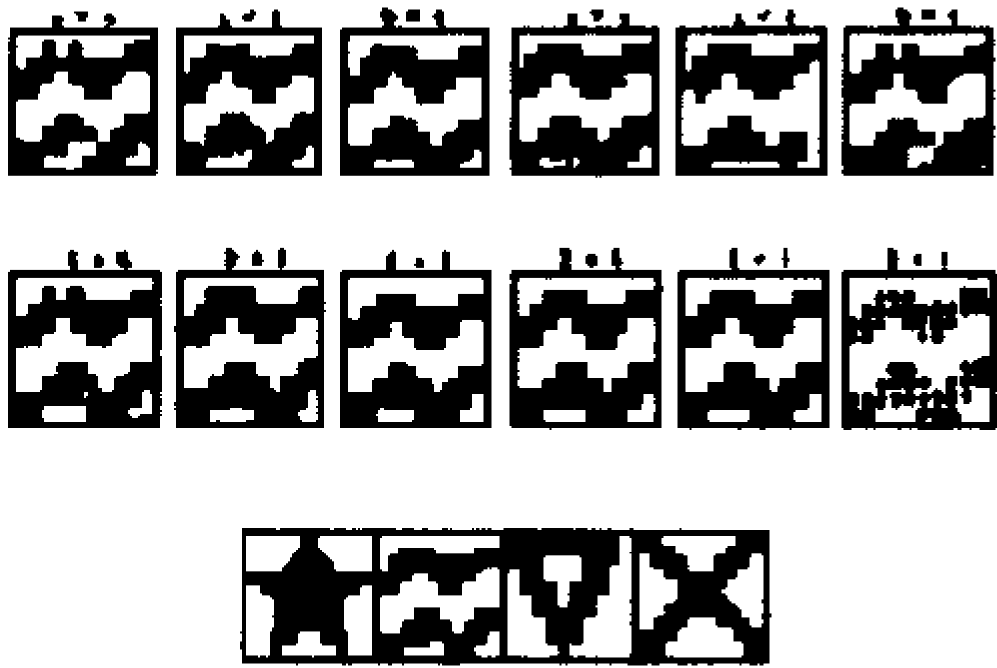


图 3.3.2

当分支参数不应性系数 a 取得比较小时, 网络行为与前述的离散反馈型 Hopfield 网络的联想记忆特性近似. 当输入 4 个记忆模式或者它们的变形时 (包括不完整模式、含有噪声的模式), 经过若干次迭代, 网络都能正确地联想出 4 个记忆模式之一, 并且稳定在该记忆模式上. 本例中 $a\leqslant 0.15$ 时, 网络呈现上述特性 (见图 3.3.2 的第二行). 从 $t=0$ 初始输入一个不完整的记忆模式, $t=1,2,\cdots,11$ 为联想出的模式结果稳定在第二种模式上, 即能正确地联想出输出结果很稳定.

另一方面, 当 a 值足够大时, 不应性影响加强, 使网络的演变过程具有非周期特性, 它不稳定于任何一个记忆模式. 又只要能保持不应性影响具有足够强度, 通过调整 a 值, 可以改变演化过程中不同时刻的网络状态, 却不能改变网络的非周期特性. 当 $a=0.5$

时可见图 3.3.3, 此时 4 个记忆模式在整个过程中都分别出现过, 但都不稳定, 说明它们是系统的不稳定吸引子.

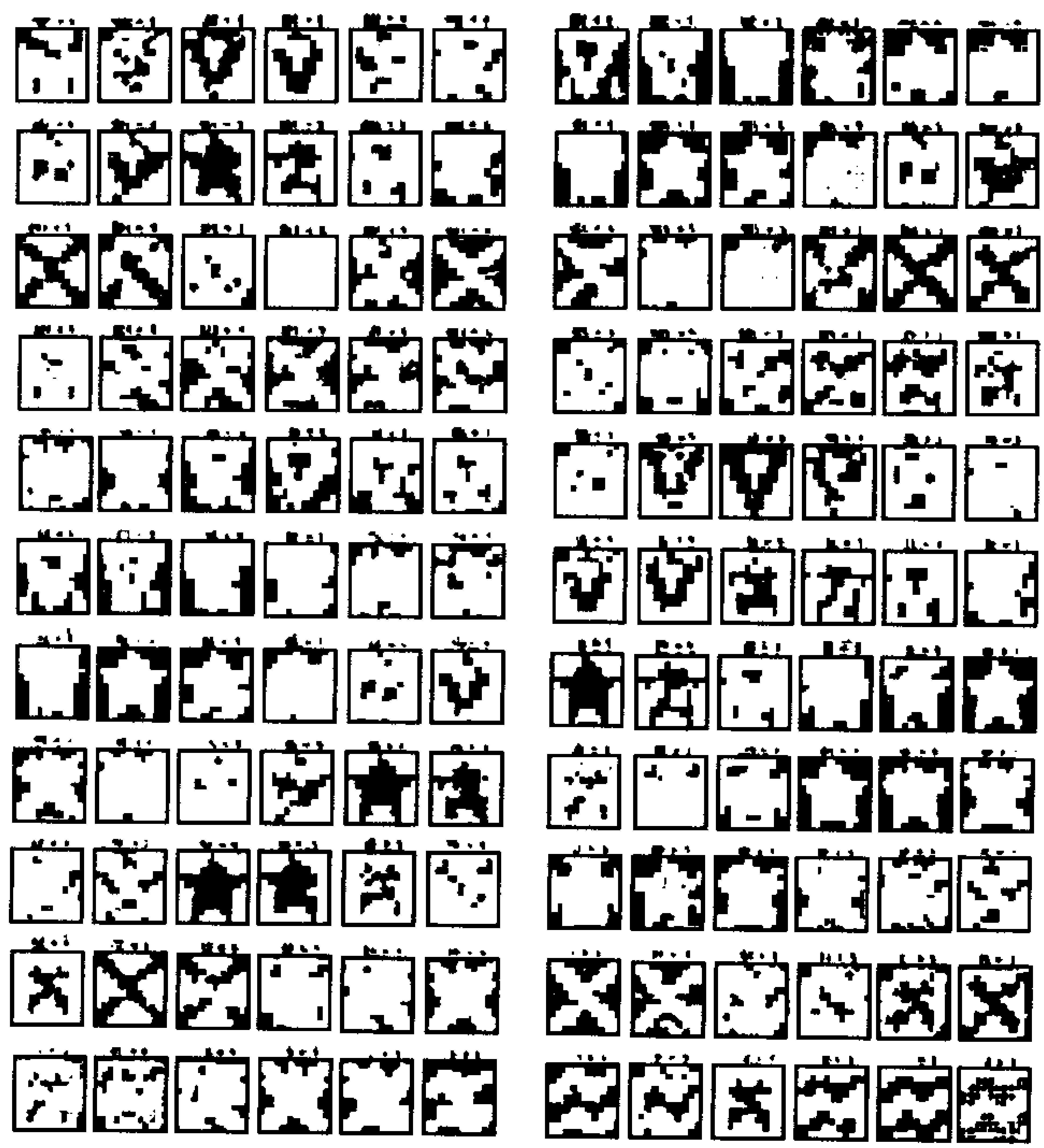


图 3.3.3　$a = 0.5$ 时网络的状态变化

模型 (3.3.13) 所示的函数图形以 x 轴为中间对称轴, 函数自变量分布于整个 y 轴, 因此网络内部状态 y_i 需用双极型表示, 即 $y_i \in (-\infty, +\infty)$. 相应地, 网络输出也必须为双极型特性, 即 $x_i \in [-1, 1]$. 当 $x_i \in [-1, 1]$ 时, 由 Hebb 规则学习得到的权值 w_{ij} 也呈双极型表示, 这些有利于网络状态的演变.

状态衰减常数 k 对网络演变也有较大影响. 严格地说, 根据 (3.3.7), k 对外部输出项、反馈项和不应项都有影响, 但在简化模型 (3.3.12) 中, k 主要影响神经元内部状态 $y_i(t)$ 的变化. 在实际模拟过程中, k 值不宜过大, 也不宜过小, 本例中选 $k = 0.8$.

由前述的动态记忆能力的分析, 可见 Aihara 混沌神经元网络

的动力学行为与以前研究过的离散 Hopfield 神经网络的静态或准静态记忆的能力完全不同, 这方面的理论与应用尚待进一步研究.

§3.4　连续的反馈型神经网络动力学模型与问题

连续的反馈型神经网络动力学模型有许多，这一节主要介绍加法模型、连接模型以及与这些模型有关的比较普遍的动力学问题 —— 收敛与吸引子问题及其分析的方法.

3.4.1　加法模型

考虑无噪声单个神经元模型，如图 3.4.1 所示.

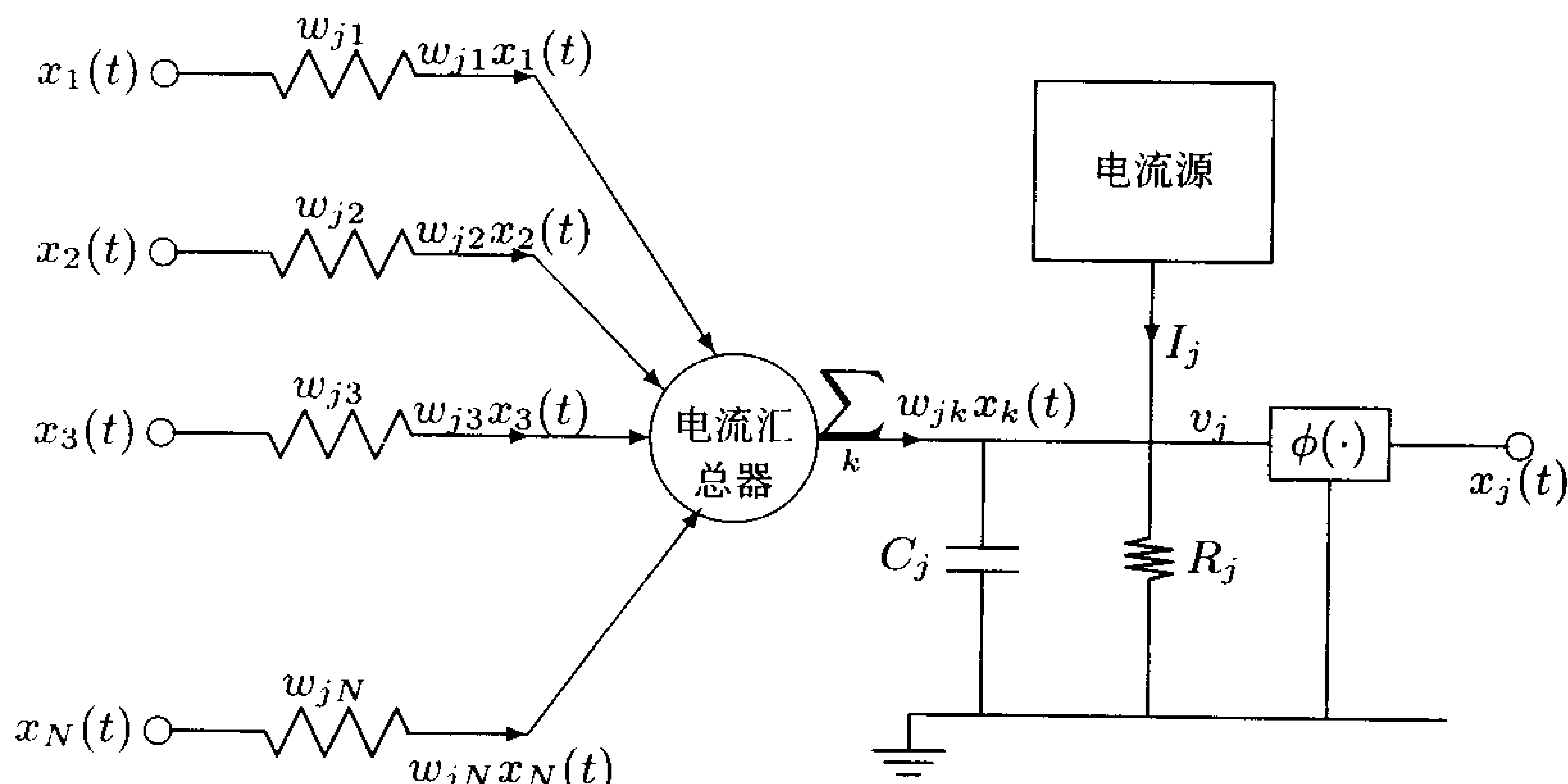


图 3.4.1　神经元的加法模型

用物理语言来说，连接权 $w_{j1},\cdots,w_{jN}$ 代表传导率，输入 $x_1(t),\cdots,x_N(t)$ 代表电势，N 是输入的个数，所有的输入均直接与电流汇合器 (current–summing junction, 简称 CSJ) 相连. 而 CSJ 具有如下 3 个特点：(i) 低输入阻抗，(ii) 综合电流的增益，(iii) 高输出阻抗. 于是 CSJ 便成为电流汇聚的结点. 在图 3.4.1 中，所有流向非线性结点的电流之和为

$$\sum_{i=1}^{N} w_{ji}x_i(t) + I_j,$$

其中第一项是由活动在权 $w_{j1},\cdots,w_{jN}$ 上的激发电位 $x_1(t),\cdots,x_N(t)$ 生成的，第二项是由电流 I_j 构成，代表外界产生的偏差. 记 $v_j(t)$ 表示通过非线性元件 $\phi(\cdot)$ 输入端的电势，于是可以得到如下所有流出非线性部分输入端的电流和:

$$\frac{v_j(t)}{R_j}+C_j\frac{\mathrm{d}v_j(t)}{\mathrm{d}t},$$

其中第一项表示非线性部分的有限阻抗，第二项与电容 C_j 有关. 于是根据 kirchoff 电流法则知，流向电路任何结点的电流和为 0，于是将之应用于图 3.4.1 的非线性元件的输入结点处，便得到

$$C_j\frac{\mathrm{d}v_j(t)}{\mathrm{d}t}+\frac{v_j(t)}{R_j}=\sum_{i=1}^{N}w_{ji}x_i(t)+I_j, \tag{3.4.1}$$

上式左端的 $C_j\dfrac{\mathrm{d}v_j(t)}{\mathrm{d}t}$ 是将动力学的内容加入神经元模型的最简单的方法. 根据电势 $v_j(t)$，可以定义神经元 j 的输出为如下非线性关系式:

$$x_j(t)=\phi(v_j(t)), \tag{3.4.2}$$

其中 $\phi(\cdot)$ 是一个非线性连续函数. 而由模型 (3.4.1) 描述的 RC 模型被称之为 **加法模型**. 用加法模型这一名称是为了区别于并联模型，在这一模型中 w_{ji} 依赖于 x_i (Grossberg, 1982).

方程 (3.4.1) 和 (3.4.2) 同样符合神经生物学的背景:

(1) $x_i(t)$ 表示第 i 个神经元命中率的短期平均值;

(2) w_{ji} 表示连接第 i 个神经元与第 j 个神经元细胞之间的有限传导率;

(3) $v_j(t)$ 表示由于输入的兴奋与抑制导致第 j 个神经元的平均电势;

(4) R_j 表示第 j 个神经元的隔膜的阻抗;

(5) C_j 表示第 j 个神经元的细胞隔膜的电容;

(6) $\phi(\cdot)$ 表示刻画第 j 个神经元输入输出特征的非线性连续函数.

现在考虑一个循环的网络，其包含 N 个相连接的神经元. 假设每一个神经元均满足方程 (3.4.1) 和 (3.4.2), 并且忽略传输过程中产生的时滞，于是可以得到一阶微分方程

$$C_j\frac{\mathrm{d}v_j(t)}{\mathrm{d}t}=-\frac{v_j(t)}{R_j}+\sum_{i=1}^{N}w_{ji}x_i(t)+I_j\quad(j=1,2,\cdots,N).\tag{3.4.3}$$

第 j 个神经元的输出 $x_j(t)$ 由一个连续可微的非线性函数 $\phi(\cdot)$ 来定义，自变量为电势 $v_j(t)$, 常用的非线性函数 $\phi(\cdot)$ 的形式为

$$\phi(v_j)=\frac{1}{1+\exp(-v_j)}\quad(j=1,2,\cdots,N).\tag{3.4.4}$$

本节中讨论的学习算法成立的一个必要条件是描述环路网络的方程组 (3.4.3) 和 (3.4.4) 具有一个不动点 (稳定的孤立的吸引子).

3.4.2 连接模型

假设对于任意的 j, 第 j 个神经元的 $R_jc_j=\tau_j$ 为一时间常数，于是对时间 t 关于时间常数 τ_j, w_{ji} 和 I_j 关于 R_j 分别正规化后，可以将 (3.4.3) 改写成

$$\frac{\mathrm{d}v_j(t)}{\mathrm{d}t}=-v_j(t)+\sum_i w_{ji}\phi(v_j(t))+I_j\quad(j=1,2,\cdots,N),\tag{3.4.5}$$

然后将方程 (3.4.2) 与之联列. 在这个联列的一阶非线性方程组中，每个神经元的电势 $v_1(t),v_2(t),\cdots,v_N(t)$ 构成了一个状态向量，它的吸引子结构与 1987 年由 Pineda 给出的下述模型的相应的结构基本上是一样的:

$$\frac{\mathrm{d}x_j}{\mathrm{d}t}=-x_j(t)+\phi\left(\sum_i w_{ji}x_i(t)\right)+k_j,\quad j=1,2,\cdots,N.\tag{3.4.6}$$

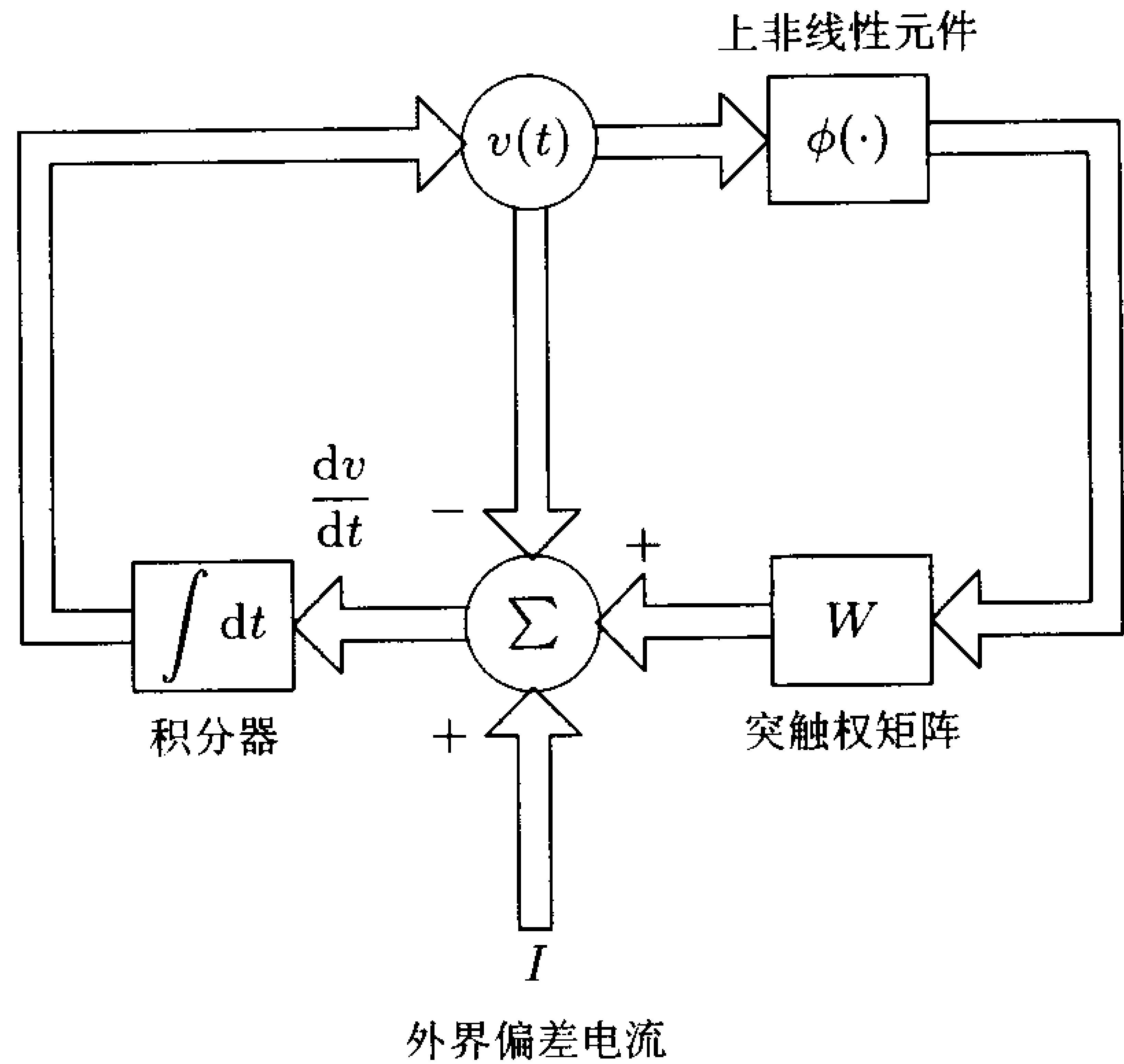
当然在上述模型中的状态变量是每个神经元的输出.

事实上，上述两类模型可以通过线性的可逆变换建立. 先在 (3.4.6) 方程两边同时乘以 w_{kj}, 并关于 j 求和，将如下变换代入:

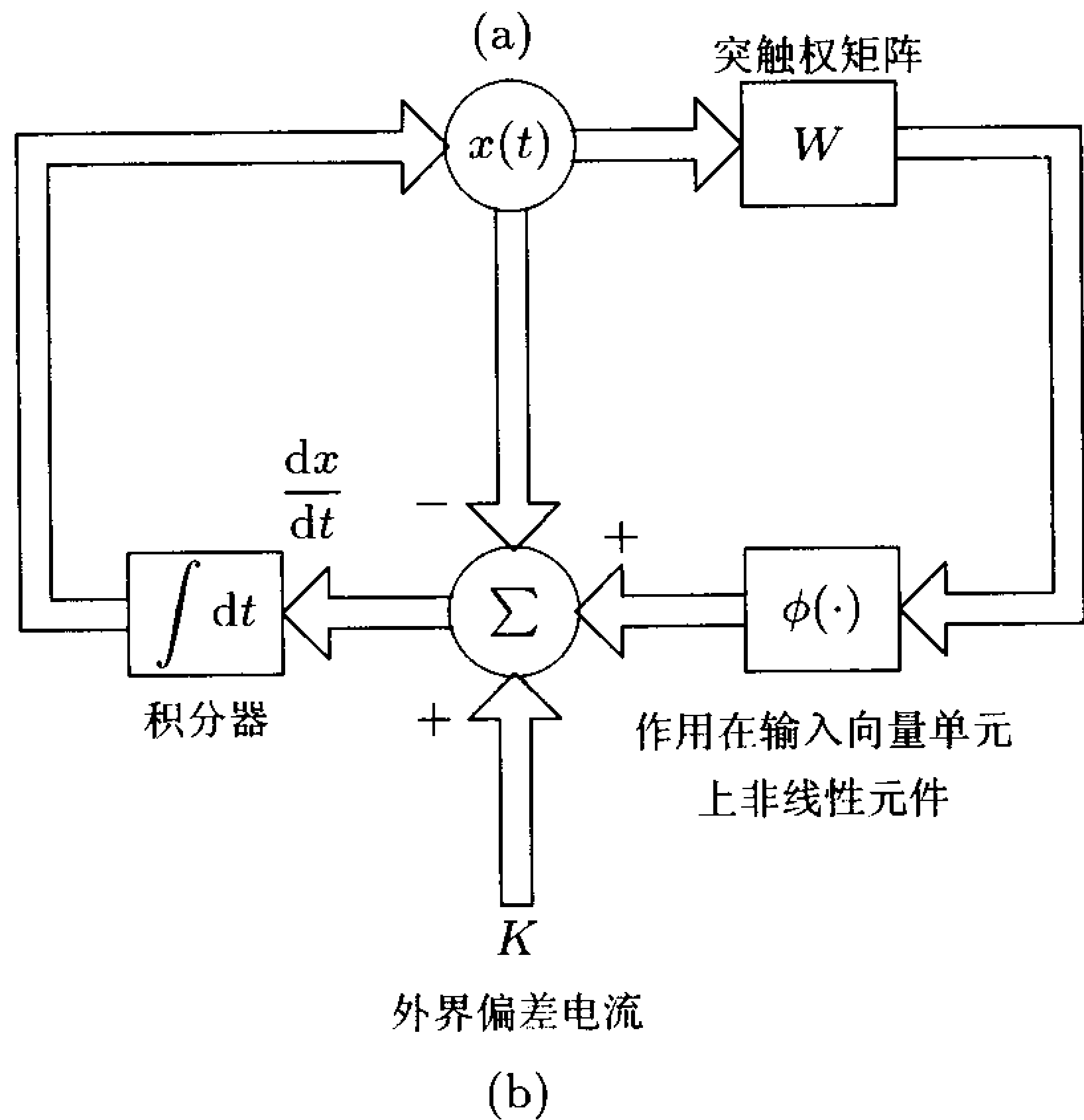
$$v_k(t)=\sum_j w_{kj}x_j(t),$$

可以得到如同 (3.4.5) 所描述的方程, 并注意到两个模型的不同项, 便有

$$I_k = \sum_j w_{kj} k_j.$$



(a)



(b)

图 3.4.2 (3.4.5) 与 (3.4.6) 两种模型的神经动力学框图

这里需要着重指出的是方程 (3.4.5) 的稳定性结果适用于方程 (3.4.6).

以上两种神经动力系统之间的关系可以由图 3.4.2 给出，从图中可以明显看出反馈的存在.

3.4.3 环路网络模型的吸引子控制

假设神经动力系统中的神经元个数 N 非常大，当不考虑噪声影响时，系统 (3.4.3) 具有大量的自由度、非线性与耗散性，于是，这样的神经网络具有复杂的吸引子，因此其可以表现出非常强大的计算能力.

神经网络模型的一项最基本的功能是由计算结果对吸引子进行识别. 为了实现这一设想，必须对网络的相平面中的吸引子位置进行训练，那么，一种采取非线性方程的学习方式可以操纵吸引子的位置，将信息翻译成所期望的形式或者学习暂时感兴趣的结构，于是，在计算算法和物理机械之间建立起一种密切的联系.

神经网络的成组性质被用来在能量消耗最小意义下执行计算任务. 例如 Hopfield 网络和盒子中的脑状态网络就有这样的功效，都可以使网络能量达到最小值，它们的差别在于其不同的应用领域. Hopfield 网络是一个非常有效的联想记忆器，也可以成为一个模拟计算机以用来对组合型优化问题的解决，而盒子中的脑状态网络模型则适用于从聚问题.

另一种被用来调整吸引子位置的方法是用最速下降来求成本函数的最小值，这方面的内容可以在由 Amari 等著的有关统计神经动力系统中找到，它也是许多学习算法的基础. 例如将一个推广的反向传播发展成环路神经网络就是这种方法的成功尝试. 特别地，由 Lapedes 和 Farber 首先将反馈应用于神经网络学习的算法之中，然而这种算法并没有运用反向传播的过程以达到修正权值的目的，而回路反向传播算法的发展在 20 世纪 80 年代后期，由 Pineda, Rohwer, Forrest, Almeida 等人各自为此做出了贡献. 其之所以有所发展，是因为人们认识到了神经网络是非线性的动力系统. 在应用中，回路反向传播算法非常符合依赖于隐藏神经元使用

性的输入输出映射.

3.4.4 连续 Hopfield 神经网络模型的动力学性质

前面讨论了离散的模型, 并采用了输出值为 0 或 1 的 McCulloch–Pitts 神经元. 本节考虑连续的情形, 特别地, 我们采用 $N=4$ 的神经网络模型 (见图 3.4.3) 来分析 Hopfield 模型的动力学性质.

连续的模型比 McCulloch–Pitts 模型更加现实, 因为它具有现实神经系统的许多特征. 有了 $x_i(t)=\phi(v_i(t))$, 可以将方程 (3.4.3) 写成如下形式:

$$C_j\frac{\mathrm{d}v_j(t)}{\mathrm{d}t}=-\frac{v_j(t)}{R_j}+\sum_i^N w_{ji}\phi(v_j(t))+I_j, j=1,2,\cdots,N. \quad (3.4.7)$$

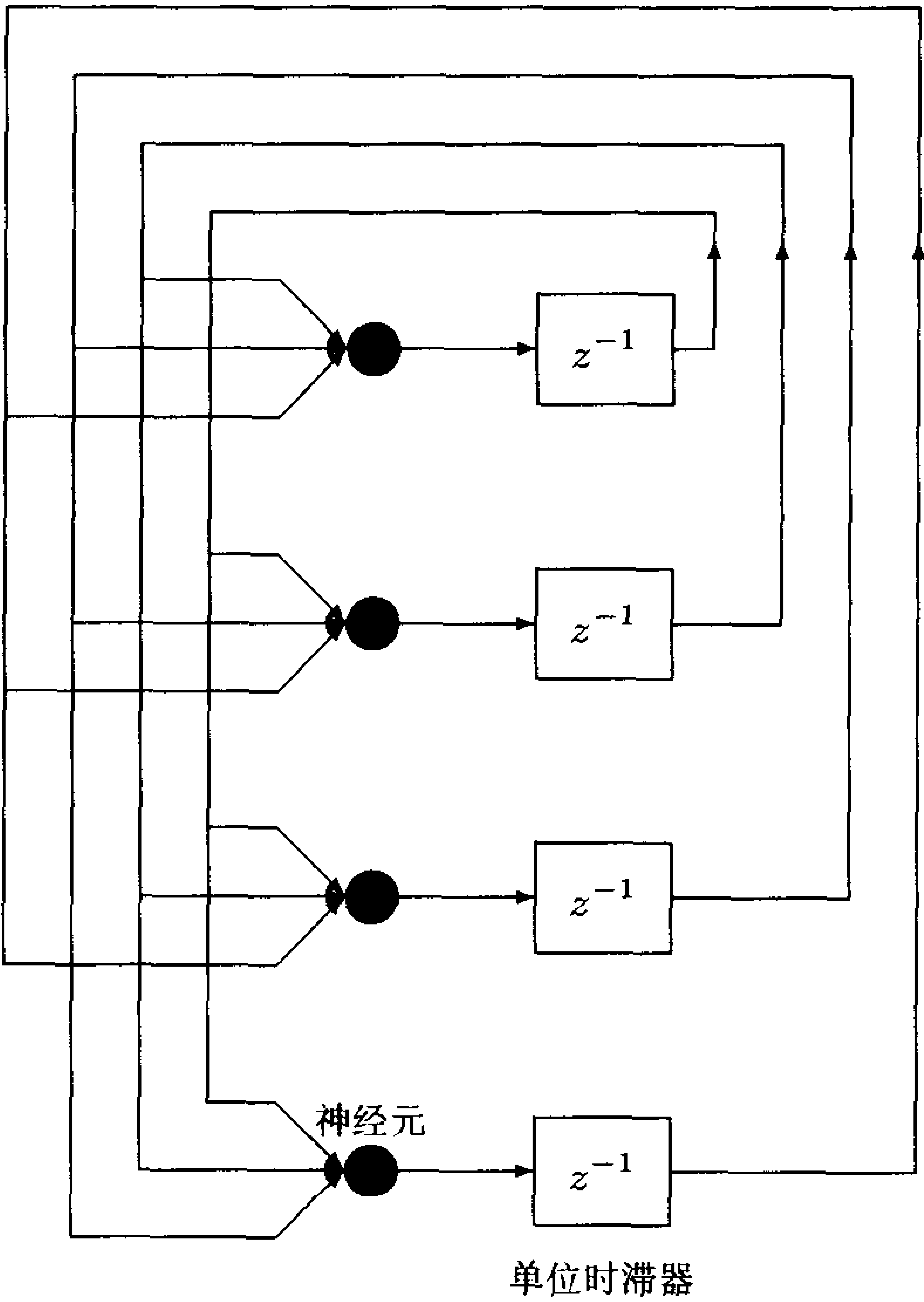


图 3.4.3

为了继续讨论，可以作如下假设：

(1) 权矩阵是对称的，即对任意 i, j 有

$$w_{ji} = w_{ij}; \tag{3.4.8}$$

(2) 任意一个神经元具有非线性性质，于是可采用非线性函数 $\phi(\cdot)$;

(3) 由函数 $\phi(\cdot)$ 的逆存在，有

$$v = \phi^{-1}(x). \tag{3.4.9}$$

令 sigmoid 函数 $\phi(\cdot)$ 为双曲正切函数

$$x = \phi_i(v) = \frac{1 - \exp(-g_i v)}{1 + \exp(-g_i v)}, \tag{3.4.10}$$

其中 g_i 表示第 i 个神经元的增益，定义为

$$g_i = \left.\frac{\mathrm{d}\phi_i}{\mathrm{d}v}\right|_{v=0}, \tag{3.4.11}$$

于是输入输出关系 (3.4.9) 的逆可以改写成如下形式：

$$v = \phi_i^{-1}(x) = -\frac{1}{g_i} \ln\left(\frac{1-x}{1+x}\right). \tag{3.4.12}$$

标准的单个神经元增益的输入输出关系逆的形式为

$$\phi^{-1}(x) = -\ln\left(\frac{1-x}{1+x}\right), \tag{3.4.13}$$

于是可以将方程 (3.4.12) 改写成

$$\phi_i^{-1}(x) = \frac{1}{g_i}\phi^{-1}(x). \tag{3.4.14}$$

图 3.4.4(a) 表示 sigmoid 函数 $\phi(v)$ 的图像，图 3.4.4(b) 表示相应的

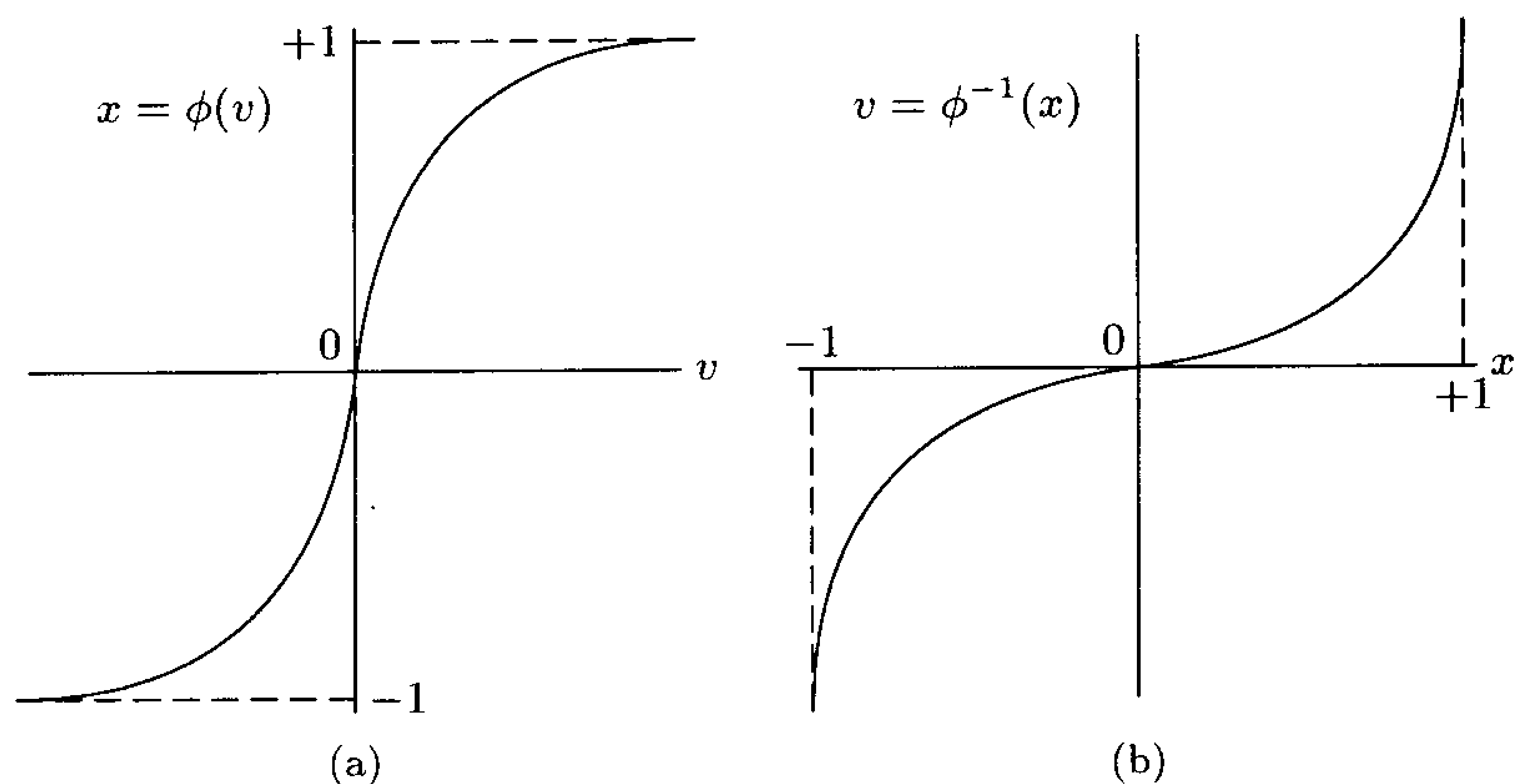


图 3.4.4　$\phi(v)$ 与 $\phi^{-1}(x)$ 的图像

非线性函数 $\phi^{-1}(x)$ 的图像.

根据 Hopfield (1984a), 环路网络的能量 (Lyapunov 函数) 可以定义为

$$E=-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}w_{ji}x_ix_j+\sum_{j=1}^{N}\frac{1}{R_j}\int_0^{x_j}\phi_j^{-1}(x)\mathrm{d}x-\sum_{j=1}^{N}I_jx_j, \tag{3.4.15}$$

对 E 关于 t 求导有

$$\frac{\mathrm{d}E}{\mathrm{d}t}=-\sum_{j=1}^{N}\left(\sum_{i=1}^{N}w_{ji}x_i-\frac{v_j}{R_j}+I_j\right)\frac{\mathrm{d}x_j}{\mathrm{d}t}. \tag{3.4.16}$$

可见上式右端括号内的就是 $C_j\dfrac{\mathrm{d}v_j}{\mathrm{d}t}$, 于是得到

$$\frac{\mathrm{d}E}{\mathrm{d}t}=-\sum_{j=1}^{N}C_j\frac{\mathrm{d}v_j}{\mathrm{d}t}\frac{\mathrm{d}x_j}{\mathrm{d}t}. \tag{3.4.17}$$

利用 v_j 与 x_j 之间的关系, 得到

$$\frac{\mathrm{d}E}{\mathrm{d}t}=-\sum_{j=1}^{N}C_j\left[\frac{\mathrm{d}}{\mathrm{d}t}\phi_j^{-1}(x_j)\right]\frac{\mathrm{d}x_j}{\mathrm{d}t}$$

$$= -\sum_{j=1}^{N} C_j \left(\frac{\mathrm{d}x_j}{\mathrm{d}t}\right)^2 \left[\frac{\mathrm{d}}{\mathrm{d}x_j}\phi_j^{-1}(x_j)\right]. \tag{3.4.18}$$

由图 3.4.4 可见函数 $\phi_j^{-1}(x_j)$ 是单调增加的，于是

$$\frac{\mathrm{d}}{\mathrm{d}x_j}\phi_j^{-1}(x_j) \geqslant 0 \quad (\text{对所有的 } i), \tag{3.4.19}$$

而且

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} \geqslant 0 \quad (\text{对所有的 } i). \tag{3.4.20}$$

综上所述，(3.4.18) 式的右端是非负的. 也就是，对于能量函数 E 有

$$\frac{\mathrm{d}E}{\mathrm{d}t} \leqslant 0 \quad (x_j \neq 0, \text{对所有的 } j). \tag{3.4.21}$$

从 E 的定义 (3.4.15) 式可知，函数 E 是有界的，又 $\frac{\mathrm{d}E}{\mathrm{d}t} = 0$ 满足的点集中的最大的不变集就是 $\frac{\mathrm{d}x_i}{\mathrm{d}t} = 0$ 满足的点集，即平衡点集. 由 Lassale 不变性原理知：由一阶非线性微分方程描述的 Hopfield 模型的演化表示在相空间的一条轨线当达到使得函数 E 取最小值的那一点后，即系统的平衡点后，停留不动. 于是可以得到如下关于 Hopfield 网络的定理

定理 3.4.1 包含 N 个神经元的 Hopfield 网络的 Lyapunov (能量) 函数 E 是一个单调递减拥有状态变量 $\{x_j \mid j = 1, 2, \cdots, N\}$ 的网络的函数，即 Hopfield 网络是全局渐近稳定的，也就是当 t 趋向于无穷大时，网络趋向于平衡点集.

Grossberg 于 1977 年第一个将 Lyapunov 函数的方法用于神经网络全局稳定性的研究分析中. 尔后，在 1983 年 Cohen 和 Grossberg 得到了一条关于方程 (3.4.7) 的 Lyapunov 函数在特殊情形下的定理，成为 Cohen–Grossberg 定理.

以下介绍另一个收敛性的结果：

定理 3.4.2 如果 W 权矩阵满足 DW 对称的条件，其中 D

是正定对角阵，那么 Hopfield 稳定收敛性定理仍然成立.

证　设模型写成

$$C\frac{\mathrm{d}x}{\mathrm{d}t}=-P(v)R+WV+I, \tag{3.4.22}$$

其中

$$C=\begin{pmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_n \end{pmatrix},\quad x=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

$$I=\begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{pmatrix},\quad R=\begin{pmatrix} 1/R_1 \\ 1/R_2 \\ \vdots \\ 1/R_n \end{pmatrix},\quad v=\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

$$P(v)=\begin{pmatrix} F^{-1}(v_1) & & & \\ & F^{-1}(v_2) & & \\ & & \ddots & \\ & & & F^{-1}(v_n) \end{pmatrix},$$

$$S(v)=\left[\int_0^{v_1}F^{-1}(\theta)\mathrm{d}\theta,\int_0^{v_2}F^{-1}(\theta)\mathrm{d}\theta,\cdots,\int_0^{v_n}F^{-1}(\theta)\mathrm{d}\theta\right]^{\mathrm{T}}.$$

因为 DW 对称，所以存在 $B\in\mathrm{R}^{n\times n}$, 使得

$$B+B^{\mathrm{T}}=DW.$$

现在定义 Lyapunov 函数

$$E = -v^{\mathrm{T}}Bv - DIV + DR^{\mathrm{T}}S(v), \tag{3.4.23}$$

则沿着方程组 (3.4.22) 的 E 的全导数为

$$\begin{aligned}
\frac{\mathrm{d}E}{\mathrm{d}t} &= -\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} Bv - v^{\mathrm{T}}B\frac{\mathrm{d}v}{\mathrm{d}t} - DI\frac{\mathrm{d}v}{\mathrm{d}t} + DR^{\mathrm{T}}P(v)\frac{\mathrm{d}v}{\mathrm{d}t} \\
&= -\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} Bv - \left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} B^{\mathrm{T}}v - \left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} I^{\mathrm{T}}D \\
&\quad + \left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} P(v)RD \\
&= -\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} [Bv + B^{\mathrm{T}}v + I^{\mathrm{T}}D - P(v)RD] \\
&= -\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} [(B + B^{\mathrm{T}})v + DI - DRP(v)] \\
&= -\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} D[Wv + I - RP(v)] \\
&= -\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} DC\frac{\mathrm{d}x}{\mathrm{d}t} \\
&= -\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^{\mathrm{T}} DCQ^{-1}\frac{\mathrm{d}v}{\mathrm{d}t} \leqslant 0.
\end{aligned}$$

这里用到 $\frac{\mathrm{d}v}{\mathrm{d}t} = Q\frac{\mathrm{d}x}{\mathrm{d}t}$, 即 $\frac{\mathrm{d}x}{\mathrm{d}t} = Q^{-1}\frac{\mathrm{d}v}{\mathrm{d}t}$, 其中

$$Q = \begin{pmatrix} F^{-1}(x_1) & & & \\ & F^{-1}(x_2) & & \\ & & \ddots & \\ & & & F^{-1}(x_n) \end{pmatrix},$$

$$Q^{-1} = \begin{pmatrix} \frac{1}{F^{-1}(x_1)} & & & \\ & \frac{1}{F^{-1}(x_2)} & & \\ & & \ddots & \\ & & & \frac{1}{F^{-1}(x_n)} \end{pmatrix}.$$

又由 $D > 0, C > 0, Q^{-1} > 0$, 知 $DCQ^{-1} > 0$.

$\frac{\mathrm{d}E}{\mathrm{d}t} = 0$ 的点集即为 $\frac{\mathrm{d}v}{\mathrm{d}t} = 0$ 的点集, 也就是平衡点集 $\frac{\mathrm{d}x}{\mathrm{d}t} = 0$. 由 Lassale 不变性原理知, (3.4.22) 的任何有界的轨线当 $t \to +\infty$ 时, 最终趋向于 (3.4.22) 的平衡点集.

当然, 对于一般的不对称的权矩阵 W, 类似 Hopfield 收敛定理仍然成立的可能性比较小, 绝大部分的动力学行为应当是振荡的或者是混沌的奇异吸引子的情形. 这方面的研究有待进一步深入.

下面介绍离散和连续的 Hopfield 模型的稳定性之间的联系.

为了给出两种情形的稳定性之间的联系, 我们重新定义单个神经元的输入和输出关系, 使之满足两条简化的特征:

(1) 神经元的输出具有渐近值

$$x_j = \begin{cases} 1, & v_j = \infty, \\ -1, & v_j = -\infty; \end{cases} \tag{3.4.24}$$

(2) 单个神经元的输入输出的中点位于原点, 表示如下:

$$\phi(0) = 0. \tag{3.4.25}$$

相应地, 置偏差 I_j 等于 0 (对于所有 j).

为了给出 Hopfield 模型连续系统的能量函数, 必须允许每个神经元有自己的回路, 而对于离散的就不需要了. 为简便起见, 置 $w_{jj} = 0$ (对于所有 j).

根据观察，重新定义连续型的能量函数

$$E=-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}w_{ji}x_ix_j+\sum_{j=1}^{N}\frac{1}{R_j}\int_0^{x_j}\phi_i^{-1}(x)\mathrm{d}x, \tag{3.4.26}$$

又由于函数 $\phi_i^{-1}(x)$ 由 (3.4.14) 定义，我们可以重新记能量函数如下：

$$E=-\frac{1}{2}\sum_{\substack{i=1\\i\neq j}}^{N}\sum_{j=1}^{N}w_{ji}x_ix_j+\sum_{j=1}^{N}\frac{1}{g_jR_j}\int_0^{x_j}\phi_i^{-1}(x)\mathrm{d}x, \tag{3.4.27}$$

而积分 $\int_0^{x_j}\phi_i^{-1}(x)\mathrm{d}x$ 的标准形式如图 3.4.5. 当 $x_j=0$, 积分值为 0, 否则积分均大于 0. 假设当 x_j 趋向于 ±1 时，积分值很大. 如果第 j 个神经元的增益 g_i 无限大，那么 (3.4.27) 式的第二项变得非常小. 特别地，当 $g_i=\infty$ 时，则连续模型的最大或最小值与离散情形的相同，于是能量函数可以简单地定义为

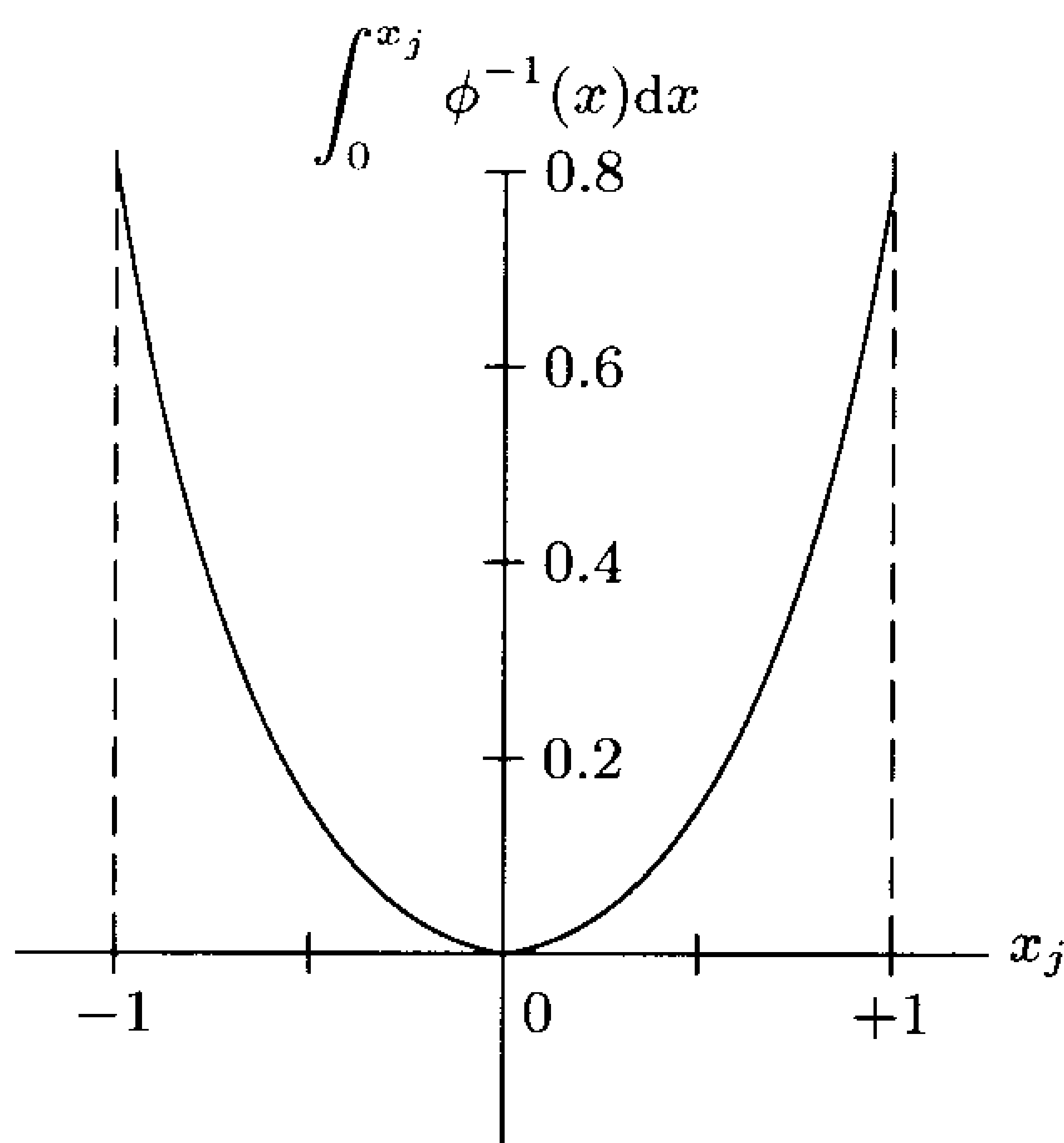


图 3.4.5　积分 $\int_0^{x_j}\phi^{-1}(x)\mathrm{d}x$ 的图像

$$E = -\frac{1}{2}\sum_{\substack{i=1\\i\neq j}}^{N}\sum_{j=1}^{N} w_{ji}x_i x_j, \tag{3.4.28}$$

其中第 j 个神经元的状态为 $x_j = \pm 1$. 我们认为，只有当增益非常大时，连续确定模型的稳定性才对应于离散的 Hopfield 模型.

然而，当每一个神经元有一个大而有限的增益 g_i 时，我们注意到 (3.4.27) 中的第二项对能量函数起了十分显著的作用. 特别地，这一作用一直能够达到单位超立方体的所有表面、边棱及顶点，而这一立方体可以用来定义模型的相空间. 另一方面，这一作用在那些远离表面的点处又相当的弱，所以，能量函数在这些顶点上有最小值，但最小值渐渐地向超立方体转移.

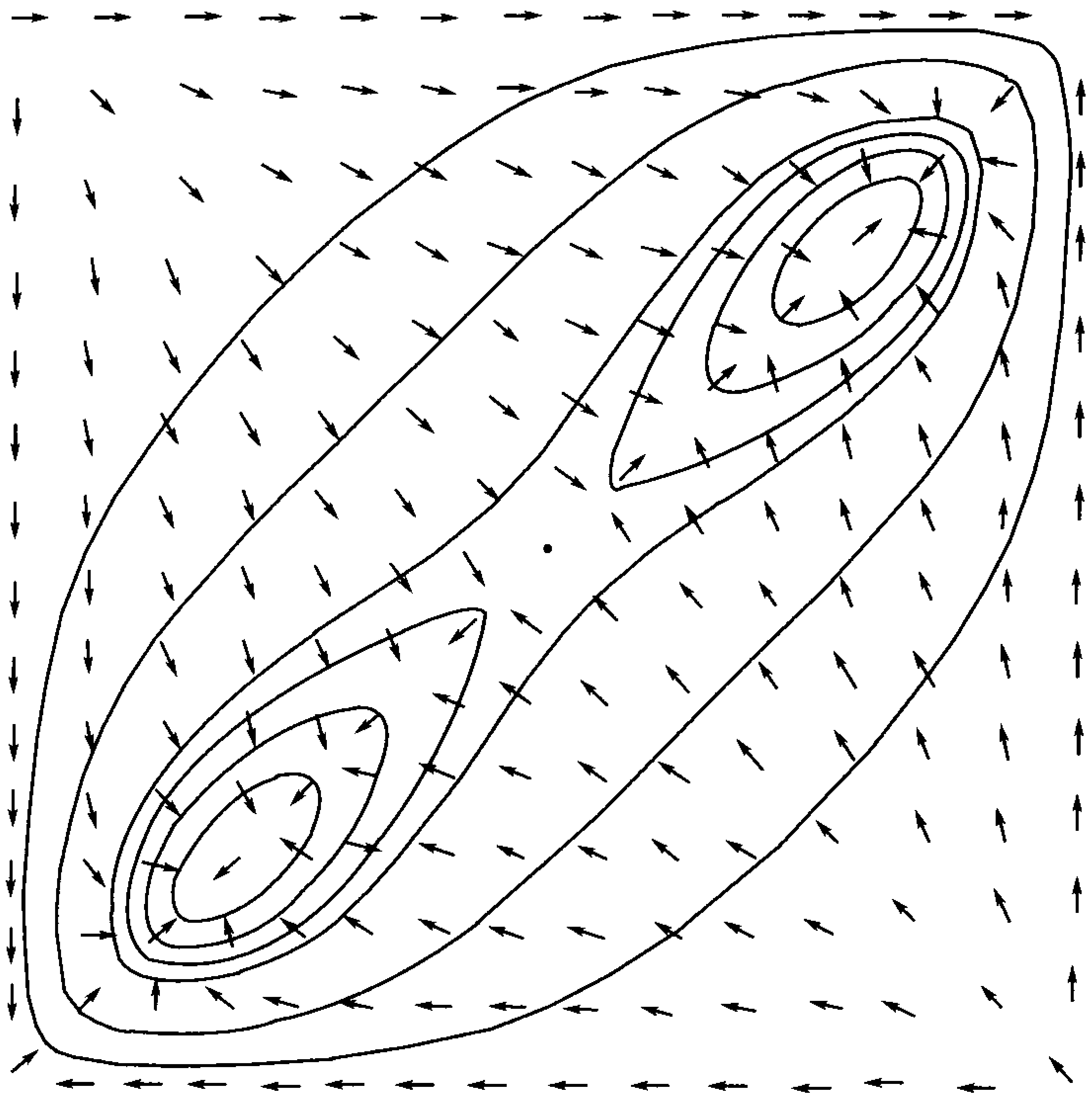


图 3.4.6

图 3.4.6 描述了有两个神经元的 Hopfield 模型的能量等高线图. 两个神经元的输出定义了图的两条坐标轴, 图 3.4.6 的左下角和右上角表示在无限增益的极限情形下稳定的最小值, 而对于有限增益的情形的最小值向内转移.

3.4.5　Cohen–Grossberg 定理

在 1983 年 Cohen 和 Grossberg 给出了判别一类神经网络稳定性的一般原理, 而这类神经网络可以用如下的非线性微分方程的并联得到:

$$\frac{\mathrm{d}u_j}{\mathrm{d}t} = a_j(u_j)\left[b_j(u_j) - \sum_{i=1}^{N} c_{ji}\phi_i(u_i)\right]. \tag{3.4.29}$$

对这个模型定义的 Lyapunov 函数为

$$E = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} c_{ji}\phi_i(u_i)\phi_j(u_j) - \sum_{j=1}^{N}\int_0^{u_j} b_j(\lambda)\phi_i'(\lambda)\mathrm{d}\lambda, \tag{3.4.30}$$

其中 $\phi_i'(\lambda) = \dfrac{\mathrm{d}\phi_i(\lambda)}{\mathrm{d}\lambda}$. 为了使定义出的 (3.4.29) 有效, 我们给出了以下假设条件:

(1) 权矩阵是对称的

$$c_{ij} = c_{ji}; \tag{3.4.31}$$

(2) 函数 $a_j(u_j)$ 满足非负性

$$a_j(u_j) \geqslant 0; \tag{3.4.32}$$

(3) 非线性输入输出函数 $\phi_i(u_j)$ 满足单调性

$$\phi_i'(u_j) \geqslant 0. \tag{3.4.33}$$

于是得到 Cohen–Grossberg 定理.

定理 3.4.3　如果给定的非线性 (3.4.29) 满足上述的对称性、非负性以及单调性三条性质, 那么由 (3.4.29) 定义的 Lyapunov 函

数 E 满足条件 $\frac{\mathrm{d}E}{\mathrm{d}t} \leqslant 0$ (对任意的 $u_j \neq 0$), 且系统是全局渐近稳定的, 即当 $t \to +\infty$ 时, (3.4.29) 的任何有界轨线最终收敛于 (3.4.29) 的平衡点集.

Hopfield 定理是 Hopfield 模型在 Cohen–Grossberg 定理意义下的特殊形式.

将系统 (3.4.29) 和 (3.4.7) 作比较, 我们可以对 Hopfield 模型及 C–G 定理作如下总结:

表 3.4.1　Hopfield 模型与 C–G 定理中模型的比较

Cohen–Grossber 定理	Hopfield 模型
u_j	$C_j v_j$
$a_j(u_j)$	1
$b_j(u_j)$	$-(v_j/R_j)+I_j$
c_{ji}	$-w_{ji}$
$\phi_i(u_i)$	$\phi_i(v_i)$

由表 3.4.1 可以产生如下 Hopfield 连续模型的 Lyapunov 函数

$$E=-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}w_{ji}\phi_i(v_i)\phi_j(v_j)+\sum_{j=1}^{N}\int_0^{v_j}\left(\frac{v_j}{R_j}-I_j\right)\phi_j'(v)\mathrm{d}v, \tag{3.4.34}$$

其中非线性函数 $\phi_j(\cdot)$ 由 (3.4.10) 给出.

我们再作如下观察:

(1) $\phi_i(v_i)=x_i$;

(2) $\int_0^{v_j}\phi_j'(v)\mathrm{d}v=\int_0^{x_j}\mathrm{d}x=x_j$;

(3) $\int_0^{v_j}v\phi_j'(v)\mathrm{d}v=\int_0^{x_j}v\mathrm{d}x\int_0^{x_j}\phi_j^{-1}(x)\mathrm{d}x$.

运用上述三项于 (3.4.34) 后, 可以得到与先前推导相同的结果, 见 (3.4.15). 需注意的是尽管 $\phi_j(v)$ 必须是关于 v 是非减函数, 但是对于 Lyapunov 函数 (3.4.34), $\phi_j(v)$ 的逆未必一定存在.

3.4.6 用非单调活化的函数修正连续的 Hopfield 网络用于联想记忆器

当 Hopfield 网络成为联想记忆器时，其中有两条缺点：

(i) 网络将受到不可避免的伪稳定状态的影响，而其个数随着学习存储向量个数的增多而增多；

(ii) 网络的存在能力，即网络学习向量的个数 (且能够被重新回忆的概率达到 99%) 等于 $\dfrac{N}{2\ln N}$.

为了克服以上的限制，已有许多设想及方法给出，其中最有意义的是 Morita (1993) 的改进方法.

由 Morita 提出的修正的方法是将每个神经元函数 $\phi(\cdot)$ 由原来的 sigmoid 函数用如图 3.4.7 中的函数来替代. 该函数可以表示为

$$\phi(v)=\left(\frac{1-\exp(-gv)}{1+\exp(-gv)}\right)\left(\frac{1+k\exp(g_0(|v|-h))}{1+\exp(g_0(|v|-h))}\right), \qquad (3.4.35)$$

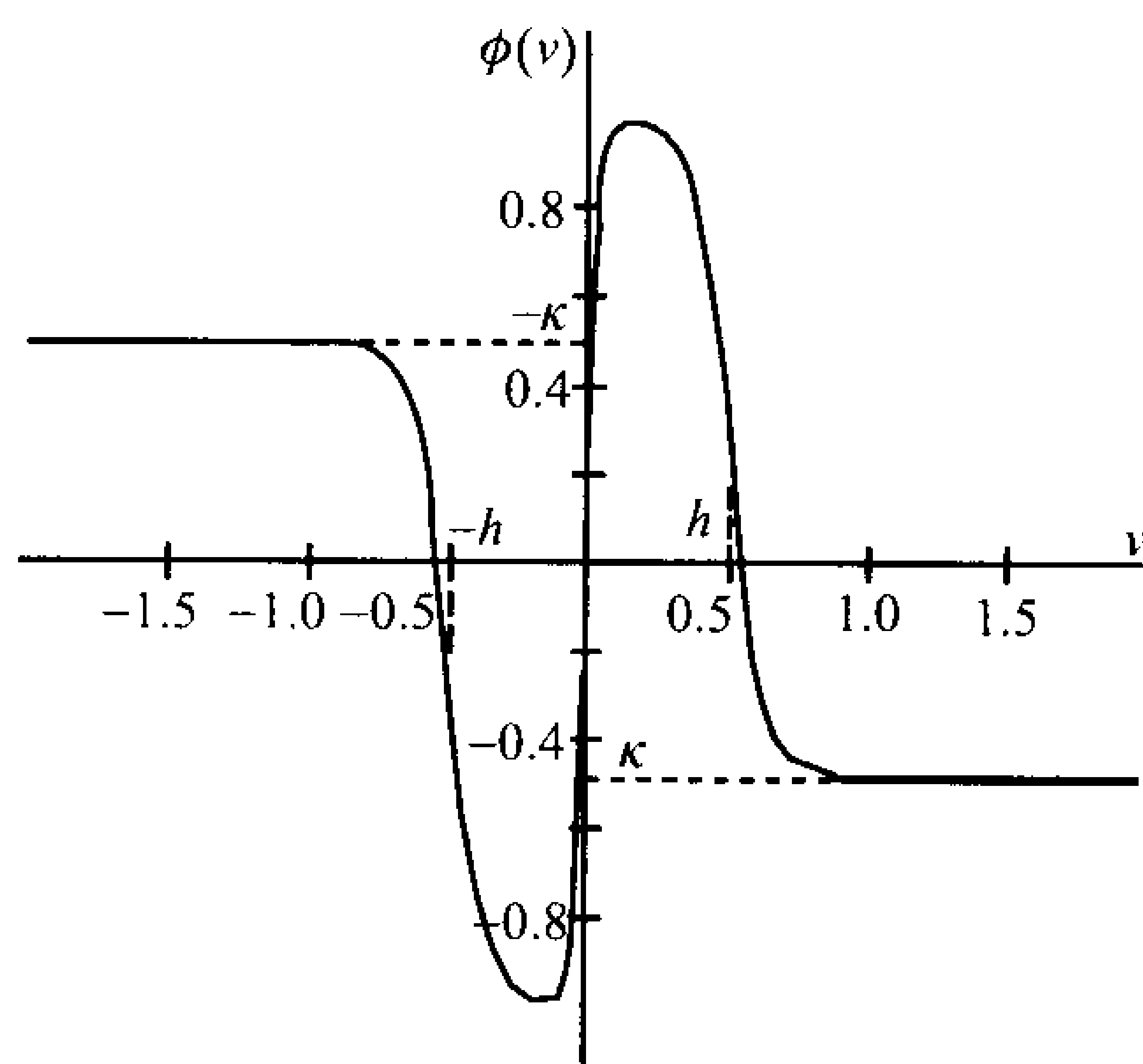


图 3.4.7　神经元的非单调的活化函数

其中 v 为活化电位，右端第一个因子是普通的 sigmoid 函数，参数 g 是在原点该因子的增益，而第二个因子的加入使得 $\phi(\cdot)$ 是非线

性的，g_0 和 h 均为常数参数，参数 k 一般为负的．由 Morita 所做的实验中，取如下参数值：

$$g = 50, \quad g_0 = 15, \quad h = 0.5, \quad k = -1.$$

按 Morita 介绍，标准的活化函数的形式为

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = -v(t) + W\phi(v(t)), \tag{3.4.36}$$

其中每个神经元的偏差设为 0, $N \times 1$ 的向量 $v(t)$ 表示神经元的活化电位，$\phi(v)$ 就是图 3.4.7 给出的函数．矩阵 W 是 $N \times N$ 的，它由学习向量 $\xi_1, \xi_2, \cdots, \xi_N$ 和下式

$$W = W_M - MI = \sum_{i=1}^{M} \xi_i \xi_i^{\mathrm{T}} - MI \tag{3.4.37}$$

给出．而回忆的过程如下，对于一个给定的试探向量 ξ_{probe}, 方程 (3.4.36) 的初值条件就取 ξ_{probe}. 如果假设解趋向于平衡点 $\overline{v}$, 则回忆模式并非取决于 $\phi(\overline{v})$, 而决定于 $\overline{v}$ 的正负号

$$\overline{x} = \mathrm{sgn}(\overline{v}).$$

如果解是不收敛的，那么整个过程是失败的，而回忆模式尚不能确定．

Morita 模型作为联想记忆有两个优点 (Yoshizana 等， 1993):

(i) 对于由 N 个神经元组成的网络，其存储量是 $0.4N$, 而相应的传统 Hopfield 网络只有 $\dfrac{N}{2\ln N}$;

(ii) Morita 模型不存在伪稳定状态，相反的当模型不能回忆正确的记忆模式，则网络的状态趋向于混沌．

Yoshizana 等 (1993) 给出了 Morita 模型理论上的分析，并用图 3.4.7 的逐段逼近函数作相应的数值模拟，很好地说明了它们的理论．

§3.5 BSB 模型、 BP 模型和环路 BP 模型及其它一些离散网络模型的动力学行为

3.5.1 BSB (Brain State in a Box) 模型及其动力学性质

1. BSB 模型

在这一节中, 我们继续讨论分析神经动力系统的记忆功能, 研究对象是由 Anderson 首先提出的 BSB 模型 (1977). BSB 模型是振幅有限的正反馈系统, 它包含连接错综复杂的神经元, 而这些神经元又以自身构成反馈. 这个模型的运作是通过运用内在的正反馈以扩大一个输出模式, 直至所有模型中的所有神经元趋向于饱和. 于是 BSB 模型被公认为是一种类型设置, 因为相近的输入模式可以由模型稳定状态定义的数值表示给出.

设 W 是对称矩阵, 且其最大的特征值有正的实部. 记 $x(0)$ 是模型的初始状态向量, 表示一个启动输入. 假设模型由 N 个神经元构成, 状态向量的维数也为 N, 矩阵 W 是 $N\times N$ 阵, 于是 BSB 模型的算法可以由以下方程完全给出:

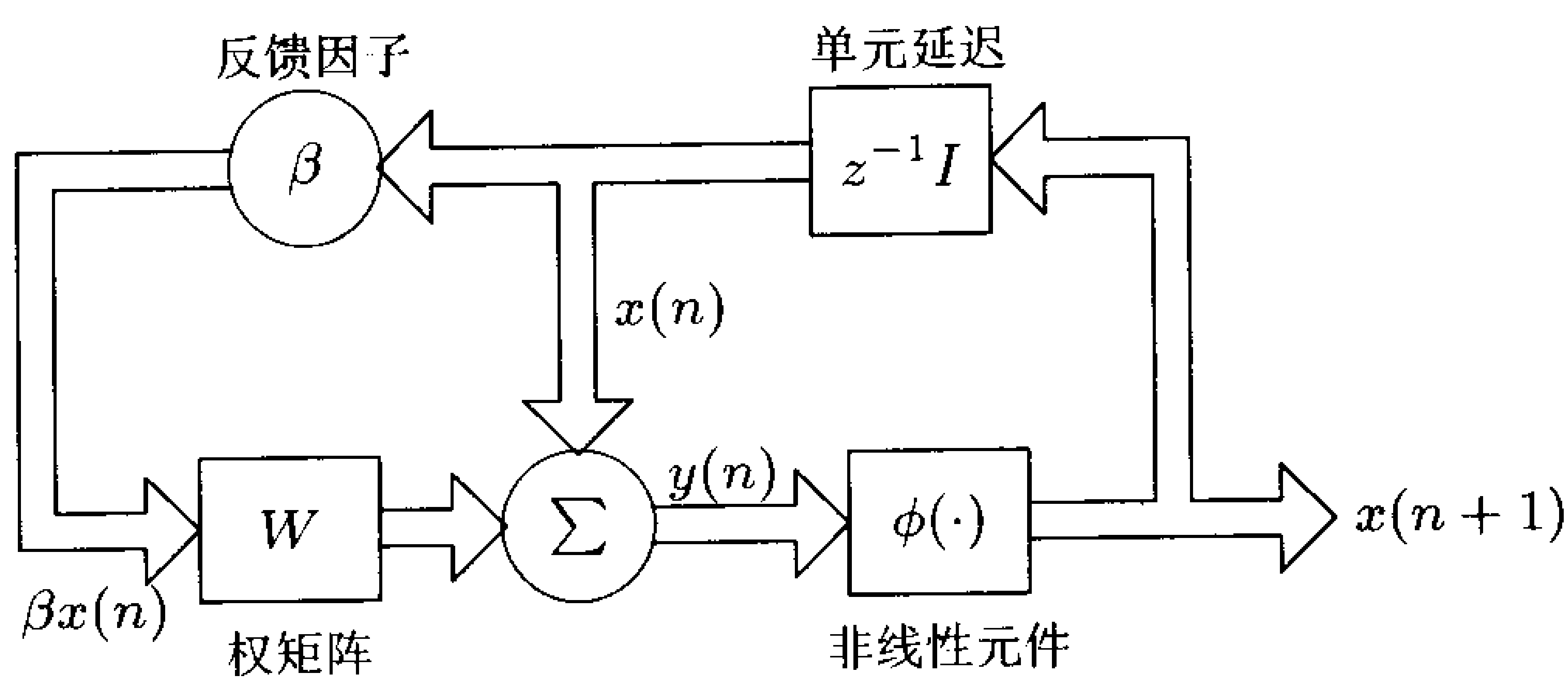
$$y(n)=x(n)+\beta Wx(n), \tag{3.5.1}$$

$$x(n+1)=\phi(y(n)), \tag{3.5.2}$$

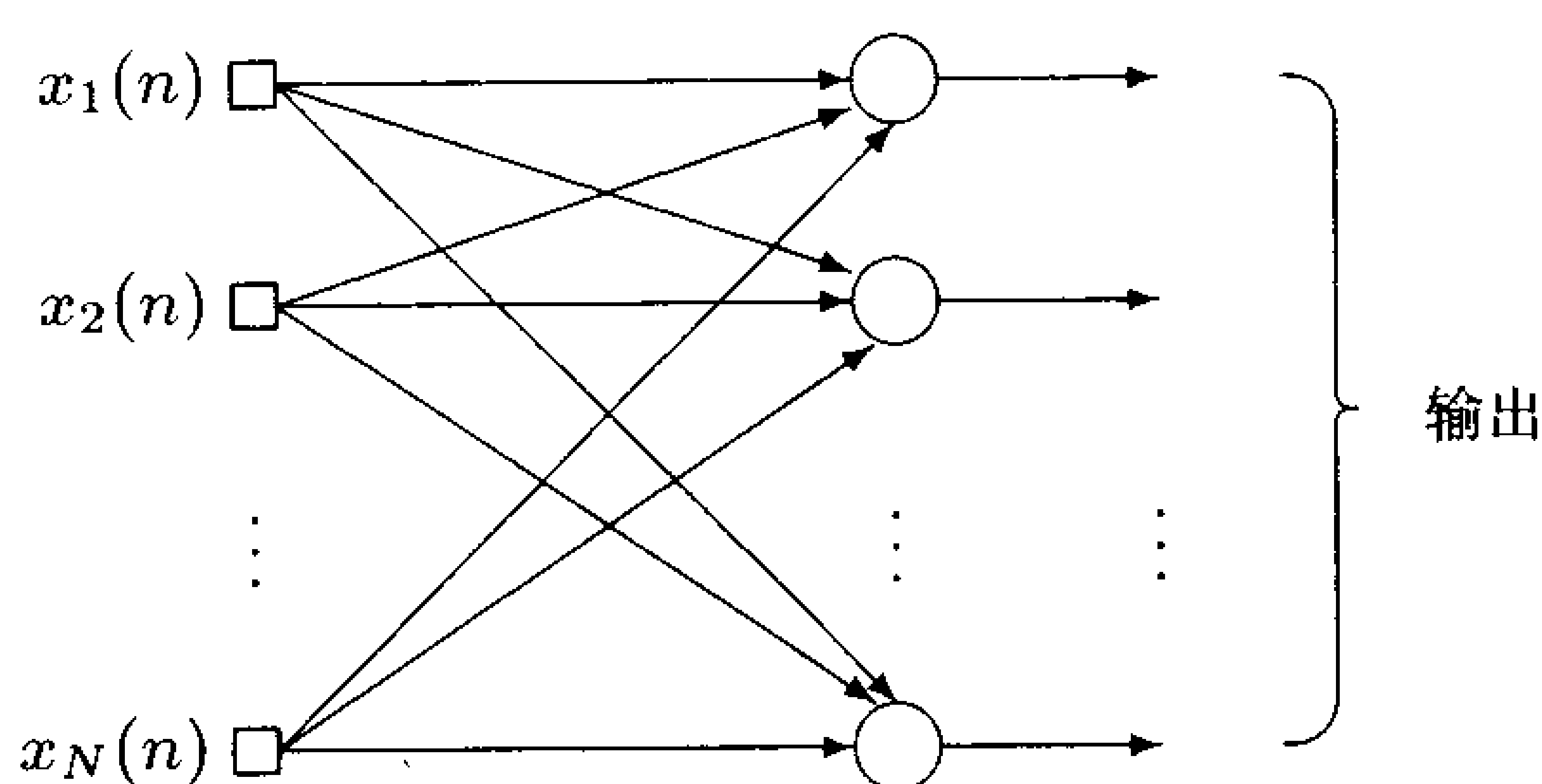
其中 β 是小正常数, 称之为**反馈因子**, $x(n)$ 是在离散时间 n 的状态向量. 图 3.5.1(a) 为方程 (3.5.1) 和 (3.5.2) 相结合的块图示, W 表示单层线性神经网络, 这由图 3.5.1(b) 给出. 函数 ϕ 是一个分段线性函数依赖于 $y_j(n)$, 其中 $y_j(n)$ 是 $y(n)$ 的第 j 个分量 (如图 3.5.2 给出).

在 $y_i(n)$ 上算法如下所述

$$x_j(n+1)=\phi(y_i(n))=\begin{cases}+1, & y_i(n)>+1,\\ y_i(n), & -1\leqslant y_i(n)\leqslant +1,\\ -1, & y_i(n)<-1.\end{cases} \tag{3.5.3}$$



(a)



(b)

图 3.5.1

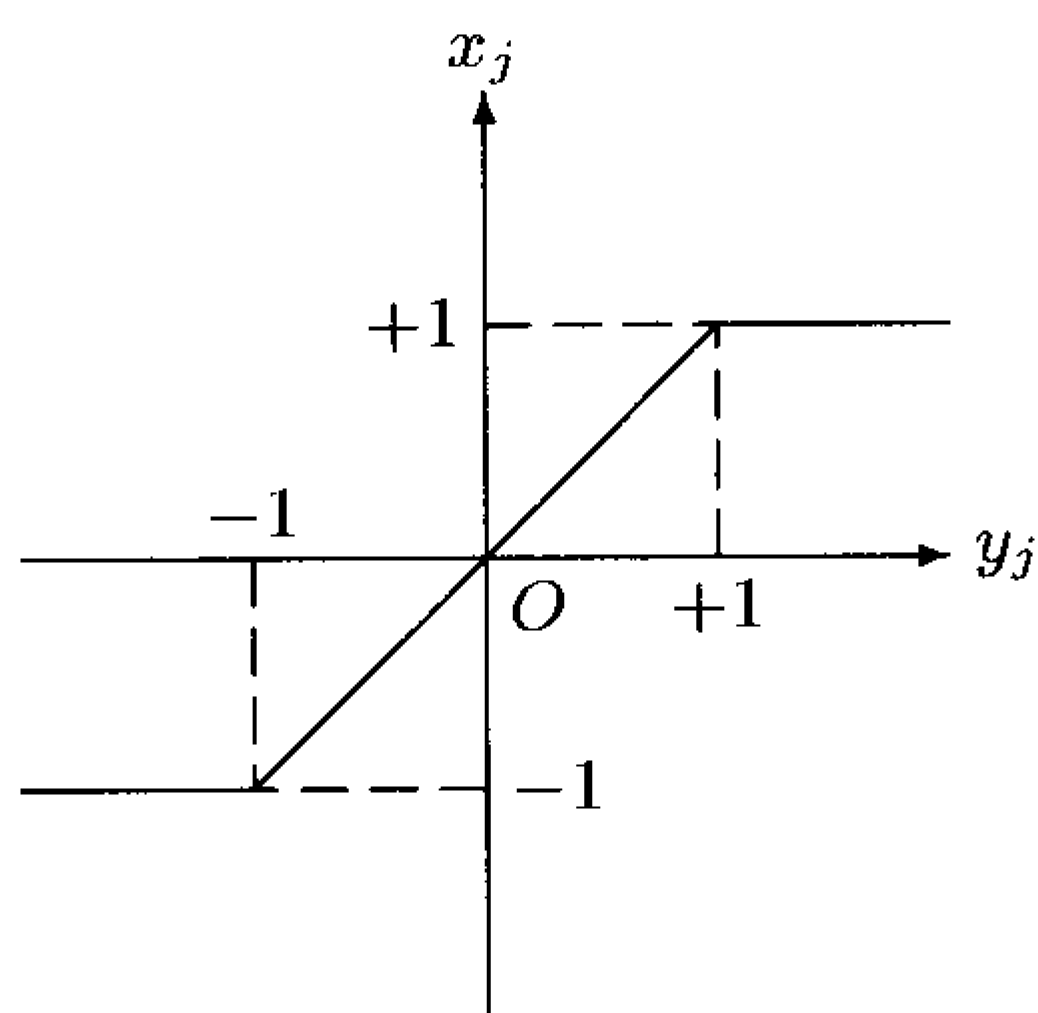


图 3.5.2

一个启动模式 $x(0)$ 作为 BSB 模型的初始状态，利用 (3.5.1) 式计算向量 $y(0)$，利用 (3.5.2) 式对 $y(0)$ 截尾，得到改进的状态向量

$x(1)$, 然后用同样的方法可以得到 $x(2)$. 这一过程直至 BSB 模型达到稳定状态为止, 所谓稳定状态即为单位超立方体的一个特别的顶点. 不难发现, 正反馈在 BSB 模型中使得状态向量 $x(0)$ 在 Euclid 距离意义下增大, 随着迭代次数的增加直至状态向量达到盒子壁, 然后沿着壁下滑, 最终到达盒子的顶点处. 在那里, 虽然状态向量还有不断地向外移动的趋势, 但仍然不会离开盒子. 这一过程由图 3.5.3 给出 ($N = 2$).

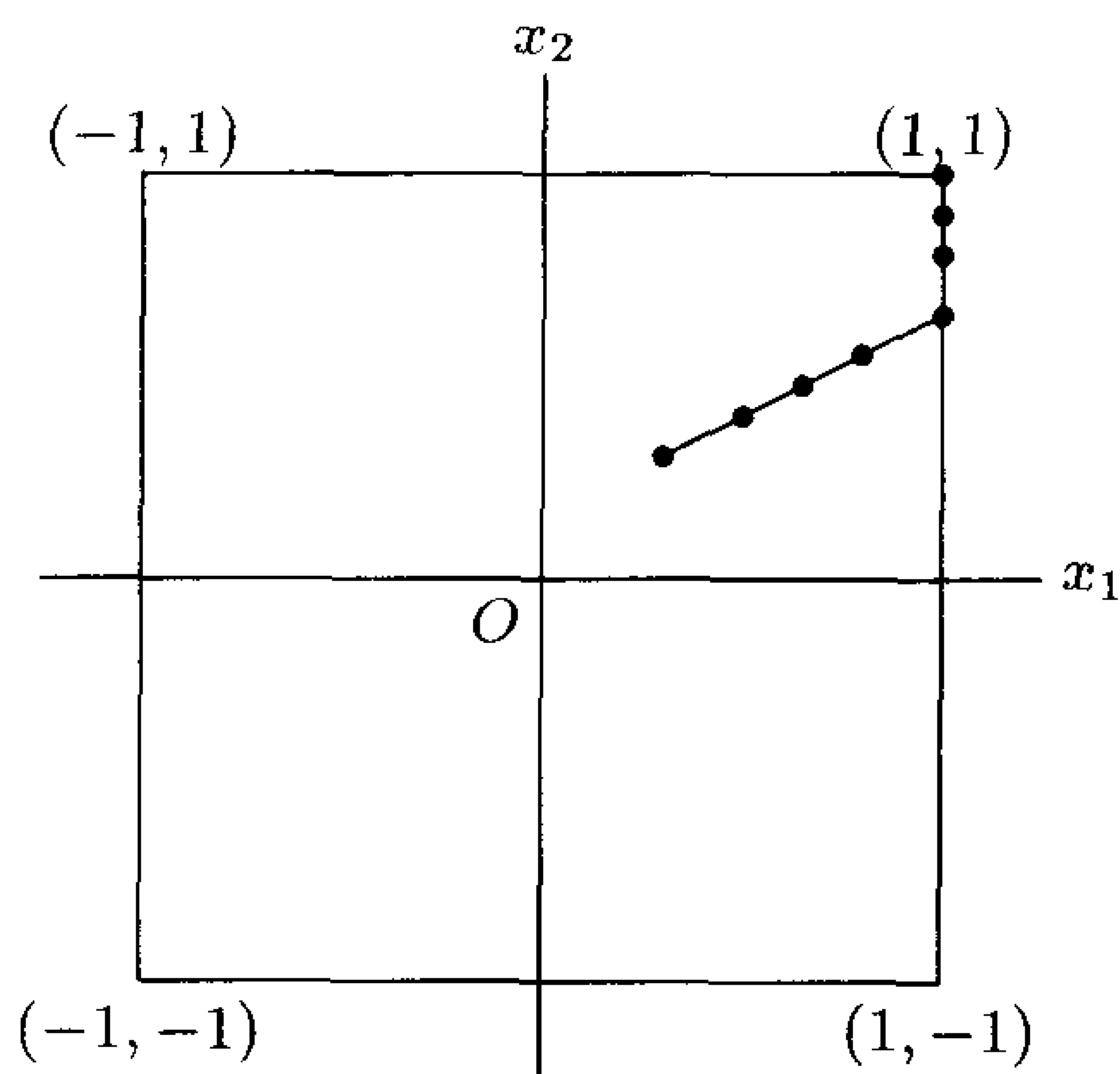


图 3.5.3

2. BSB 模型的 Lyapunov 函数及动力学性质

BSB 模型可以重新定义 (Grossberg, 1990). 为了说明这一点, 首先将 BSB 模型的算法 (3.5.1) 和 (3.5.2) 重新改写第 j 个部分, 形式如下:

$$x_j(n+1) = \phi\left(\sum_{i=1}^{N} c_{ji}x_i(n)\right), \quad j = 1, 2, \cdots, N, \tag{3.5.4}$$

其中系数 c_{ji} 定义如下:

$$c_{ji} = \delta_{ji} + \beta w_{ji}, \tag{3.5.5}$$

其中 δ 是 Kronecker 记号，且 w_{ji} 是矩阵 W 的元素. 方程 (3.5.4) 以离散的形式写成，为了继续研究，需要将之写为连续的形式

$$\frac{\mathrm{d}x_j(t)}{\mathrm{d}t} = -x_j(t) + \phi\left(\sum_{i=1}^{N} c_{ji}x_i(t)\right), \quad j = 1, 2, \cdots, N, \tag{3.5.6}$$

其中偏差 $I_j = 0$. 为了应用 C–G 定理, 我们还需更进一步将 (3.5.6) 式转换成类似于加法模型的形式. 我们可以通过引进一组新的变量

$$v_j(t) = \sum_{i=1}^{N} c_{ji}x_i(t) \tag{3.5.7}$$

来做到这一点.

相应地可以将 (3.5.6) 式改写成等价的形式

$$\frac{\mathrm{d}v_j(t)}{\mathrm{d}t} = -v_j(t) + \sum_{i=1}^{N} c_{ji}\phi(v_i(t)), \quad j = 1, 2, \cdots, N. \tag{3.5.8}$$

现在可以应用 C–G 定理了. 将 BSB 模型及 C–G 定理中相应的变量列于表 3.5.1 中，于是可以构成 BSB 模型的 Lyapunov 函数

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} c_{ji}\phi(v_j)\phi(v_i) + \sum_{j=1}^{N}\int_0^{v_j} v\phi'(v)\mathrm{d}v, \tag{3.5.9}$$

其中 $\phi'(v)$ 是 sigmoid 函数 $\phi(v)$ 关于 v 的一阶导数. 最后将 (3.5.3), (3.5.5), (3.5.7) 代入

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} = -v(t) + w\phi(v(t)) \tag{3.5.10}$$

中，定义 BSB 模型的最初变量构成的 Lyapunov 函数

$$E = -\frac{\beta}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ji}x_jx_i = -\frac{\beta}{2}x^{\mathrm{T}}Wx. \tag{3.5.11}$$

表 3.5.1

Cohen–Grossber 定理	BSB 模型
u_j	$C_j v_j$
$a_j(u_j)$	1
$b_j(u_j)$	$-v_j$
c_{ji}	$-c_{ji}$
$\phi_j(u_j)$	$\phi_j(v_j)$

对于 Hopfield 网络的 Lyapunov 函数的计算在前面已经给出. 在那里假设非线性 sigmoid 函数逆的一阶导数存在, 这样便可以应用双曲正切函数. 然而, 在 BSB 模型中, 这一条件在神经元取值为 1 或 -1 时并不满足. 尽管有此困难, BSB 模型的 Lyapunov 函数仍可以通过 C–G 定理的计算得到, 这一点在介绍该定理时已经详细说明了.

直接的分析由 Golden (1986) 给出, 他证明了 BSB 模型事实上是使能量函数 E 达到最小值的梯度下降法. 这一性质的成立是基于以下两个条件:

(1) 矩阵 W 为对称的: $W = W^{\mathrm{T}}$.

(2) W 为半正定阵.

随着 n 的增加, 能量函数 E 逐渐减少, 而且在 $n+1$ 时刻的状态 $x(n+1)$ 与 $x(n)$ 不同. E 的最小值点为 BSB 模型的不动点状态, 即 $x(n+1) = x(n)$, 也就是, BSB 模型像 Hopfield 网络一样皆为能量最小网络.

BSB 模型的不动点由单位超立方体的顶点及原点定义. 在后一种情形, 任何对状态向量的扰动, 都会被正反馈放大, 从而使模型的状态离开原点向稳定的流形移动, 也就是说原点是鞍点. 为了使每一个立方体的顶点均成为不动点, 那么权矩阵 W 必须满足第 3 个条件:

(3) 权矩阵 W 是对角占优的, 即

$$w_{jj} \geqslant \sum_{i \neq j} |w_{ij}|, \quad j = 1, 2, \cdots, N. \tag{3.5.12}$$

为了使状态 x 是稳定的, 即对一个特定的顶点是一个吸引子, 那么存在一个吸引域 $N(x)$ 位于立方体中, 使得所有初值位于 $N(x)$ 中的 $x(0)$ 的状态收敛于 x. 为了使每一个顶点均成为吸引子, 权矩阵 W 必须满足

(4) W 是严格对角占优的, 即

$$w_{jj} \geqslant \sum_{i \neq j} |w_{ij}| + \alpha, \quad j = 1, 2, \cdots, N, \tag{3.5.13}$$

其中 α 是正常数.

这里必须指出的是, 矩阵 W 为对称半正定的只能保证一部分顶点为吸引子. 为了使所有的顶点成为吸引子, 那么矩阵 W 必须满足 (3.5.13), 当然首先应当满足 (3.5.12).

3. BSB 模型的应用 —— 分类

BSB 模型的一项基本的应用是分类, 这是由于作为吸引子的超立方体的顶点和相应的吸引域将相空间划分成相应的定义完善的区域, 于是 BSB 模型被用来作为一种无监护的分类算法, 而每一个不动点代表一组相关的数据构成的类, 由正反馈产生的自我放大是分类方法的最重要的特性之一.

Anderson 于 1990 年描述了利用 BSB 模型进行分类的方法, 它可以从不同的放射射线中区分出辐射的信号. 在这一应用中, 权矩阵是通过误差纠正的线性相关的学习过程得到的. 为了详细说明, 假设有 K 个训练的向量

$$x_k \to x_k \quad (k = 1, 2, \cdots, K). \tag{3.5.14}$$

随机选择向量 x_k, 于是权矩阵可以根据偏差纠正算法得到

$$\Delta W = \eta(x_k - W x_k) x_k, \tag{3.5.15}$$

其中 η 是学习率参数. 根据计算, 输入向量 x_k 像其 “老师” 一样表现, 因为线性结合会重新构造输入向量, 而学习的目标是使得

$$W x_k = x_k, \quad k = 1, 2, \cdots, K. \tag{3.5.16}$$

由式 (3.5.14) 给出的纠偏算法在最小均方意义下接近于理想的条件 (3.5.15). 学习过程是为了促使线性结合器给出一组特征向量, 相应的特征值等于 1.

为了模拟雷达分类, BSB 模型应用上述方法给出了权矩阵. 具体的计算方式如下:

$$x(n+1)=\phi(\gamma x(n)+\beta Wx(n)+\delta x(0)), \tag{3.5.17}$$

这与 (3.5.1) 和 (3.5.2) 两式给出的算法有两点不同:

(i) 衰变项 γ 使得当前的状态发生衰变, 如果假设 γ 是一个小于 1 的正常数, 那么误差最终将趋向于 0;

(ii) 第三项 $\delta x(0)$ 是为了使初始状态 $x(0)$ 始终存在, 它具有对 BSB 模型的状态影响.

不断的迭代使得具有大的正特征值的权矩阵 W 的特征向量占优, 于是向量 $x_1,\cdots,x_K$ 被线性结合器学习. BSB 模型的分类能力通过模型中的正反馈被加强了, 所以经过若干次迭代后, 其控制了系统的状态. 另一方面, 与噪声相关的特征向量总与小特征值联系在一起, 所以为了减少对 BSB 模型状态的影响, 可以假设收到的信号相应于噪声而言是非常强的.

3.5.2 BP (Back–Propagation) 模型

BP 网络由非线性变换单元组成的多层的前馈网络, 其中无反馈.

1. 网络的结构及数学模型

BP 网络是多层网络, 有输入层、输出层及中间隐层. 最简单的就是中间隐层仅有一层的情形, 即共计有 3 层的网络. 每层具有的神经元个数可以不一样, 前一层的神经元与后一层的神经元之间有权连接, 且每个神经元的非线性活化函数最简单的就是 S 函数, 例如取为 $\phi(u)=\dfrac{1}{1+\mathrm{e}^{-u}}$.

举 1 个 3 层的网络为例 (如图 3.5.4)

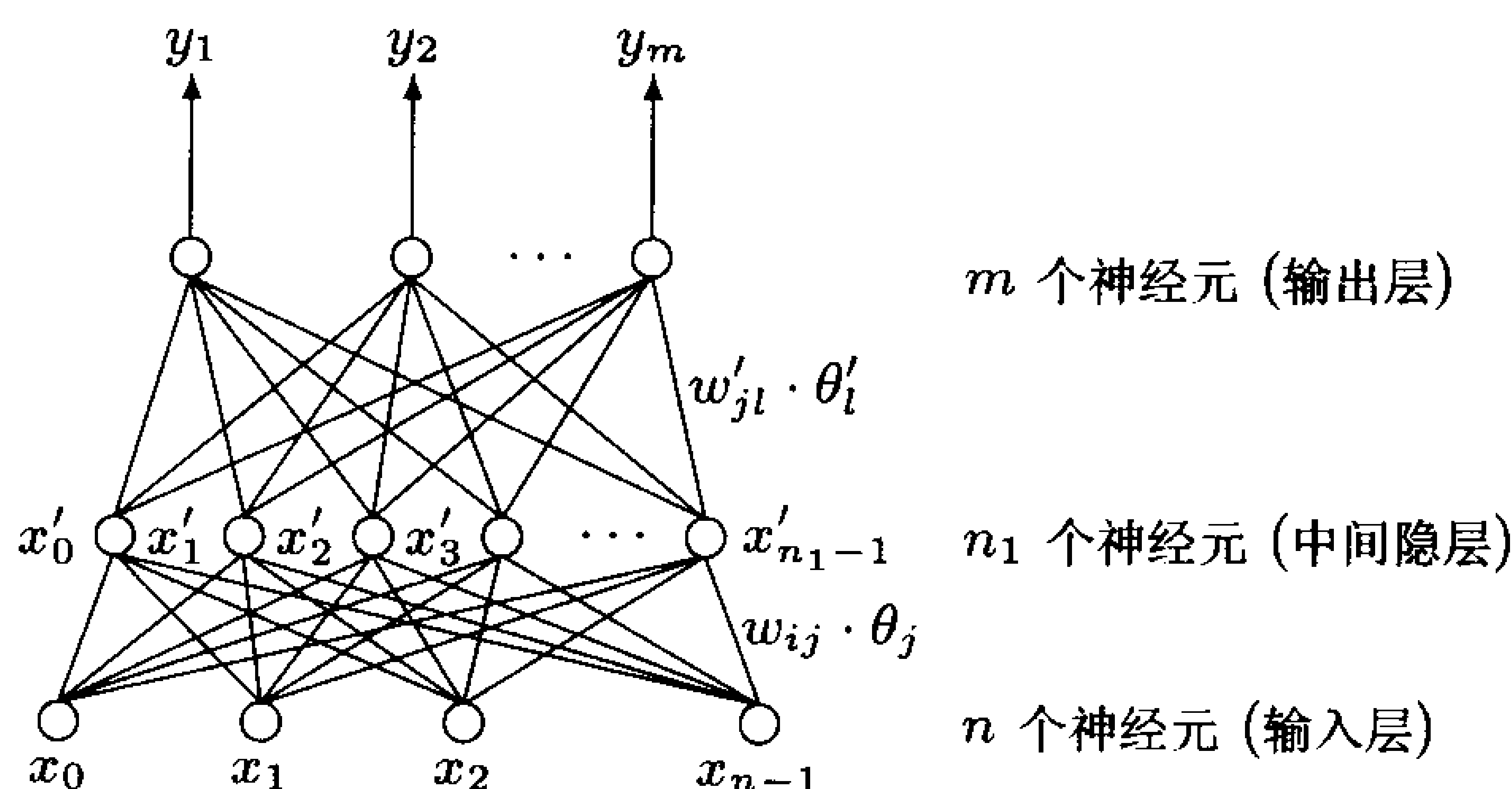


图 3.5.4　3 层 BP 网络

每一层的输入输出的关系为

$$\begin{cases} y_l = \phi\left(\sum_{j=0}^{n_1-1} w'_{jl}x'_j - \theta_l\right), & l = 1,2,\cdots,m, \\ x'_j = \phi\left(\sum_{j=0}^{n-1} w'_{ij}x'_i - \theta_j\right), & j = 1,2,\cdots,n_1. \end{cases}$$

对于该网络, 若第 P_1 个样本 $x^{P_1} = (x_0^{P_1}, x_1^{P_1}, \cdots, x_{n-1}^{P_1})$ 为输入, 则有输出 $y^{P_1} = (y_0^{P_1}, y_1^{P_1}, \cdots, y_{n_1-1}^{P_1})$. 设有 P 个学习样本 $x^1, x^2, \cdots, x^P$, 对应的"教师"样本也有 P 个 $t^1, t^2, \cdots, t^P$, 注意对每个 P_1, 有 $t^{P_1} = (t_0^{P_1}, t_1^{P_1}, \cdots, t_{n_1-1}^{P_1})$, $P_1 \in \{1, 2, \cdots, P\}$. 数学问题是要找一个非线性的映射 $F: \mathrm{R}^n \to \mathrm{R}^m$, 使得输入 P 个样本 $x^1, x^2, \cdots, x^P$, 得到的输出恰为"教师"样本 $t^1, t^2, \cdots, t^P$. 一般说来, 该非线性的映射 F 很难求得, 也就是所谓的暗箱问题. 但是可以通过 3 层 BP 网络或多层 BP 网络来近似地求得该映射 F, 关键是要适当选取权矩阵及阈值向量 Θ.

可以建立一套 BP 算法使得选择的 W, Θ (实际上增加状态变量的维数, 可以把 Θ 归于 W 之中), 不断通过梯度法去修正 w_{sq}, 即 $w_{sq} := w_{sq} + \Delta w_{sq}$, 目标是使得由初始输入 $x^1, x^2, \cdots, x^P$ 得到输出 $y^1, y^2, \cdots, y^P$ 与"教师" $t^1, t^2, \cdots, t^P$ 误差越来越小.

2. BP 学习算法过程的两个阶段

一方面从网络的输入层向前计算，若网络的结构和权已经设定，输入已知的样本，则可以计算每一层的神经元输出. 另一方面从最后输出层向后倒算修改权和阈值，从已知输出层与教师的误差修改与输出层相联的权，然后倒退修改各层的权. 将这一过程的两个阶段反复交替，直到达到收敛以及误差小于设定的值为止.

这方面的算法设计步骤与应用在最后一章中将给予介绍，本节只是为了能与下一节的环路 BP 网络模型作比较，简单叙述 BP 网络的设计思路，这里网络全部是前馈式的.

3.5.3 环路 BP 模型

考虑这样的神经网络，它包括 N 个相互连接、取值连续的神经元，关于时间与振幅连续，且连接第 i 和第 j 个神经元的权重为 w_{ji}. 所有的神经元假设都有相同的非线性激化函数 $\phi(v)$, v 是启动势. 与 3.5.2 中 BP 网络模型不同的是，这里网络中的神经元有反馈存在，即网络构成环路.

3 种神经元构成的子集在神经网络中需要加以区分.

(1) 输入神经元，它们直接接收外面提供的模式;

(2) 输出神经元，它们提供全面的网络学习以后产生的反映;

(3) 隐含神经元，既非输入也非输出神经元.

注意一个神经元可以同时成为输入和输出神经元，这样的神经元被称为自动启动式的.

图 3.5.5 说明了上面所描述网络的结构. 神经元 1 和 2 为输入， 4 和 5 为输出， 3 为仅有的隐含神经元， 5 既是输出神经元也是输入神经元，所以这是自启动式的.

外界的影响是通过偏差项来影响网络的. 如果对第 j 个神经元，那么 $I_j = \xi_j$ (ξ_j 为输入模式的一个元素), 否则 $I_j = 0$, 于是输入关系可以表述为

$$I_j = \begin{cases} \xi_j, & \text{对输入神经元}, \\ 0, & \text{其它}. \end{cases} \tag{3.5.18}$$

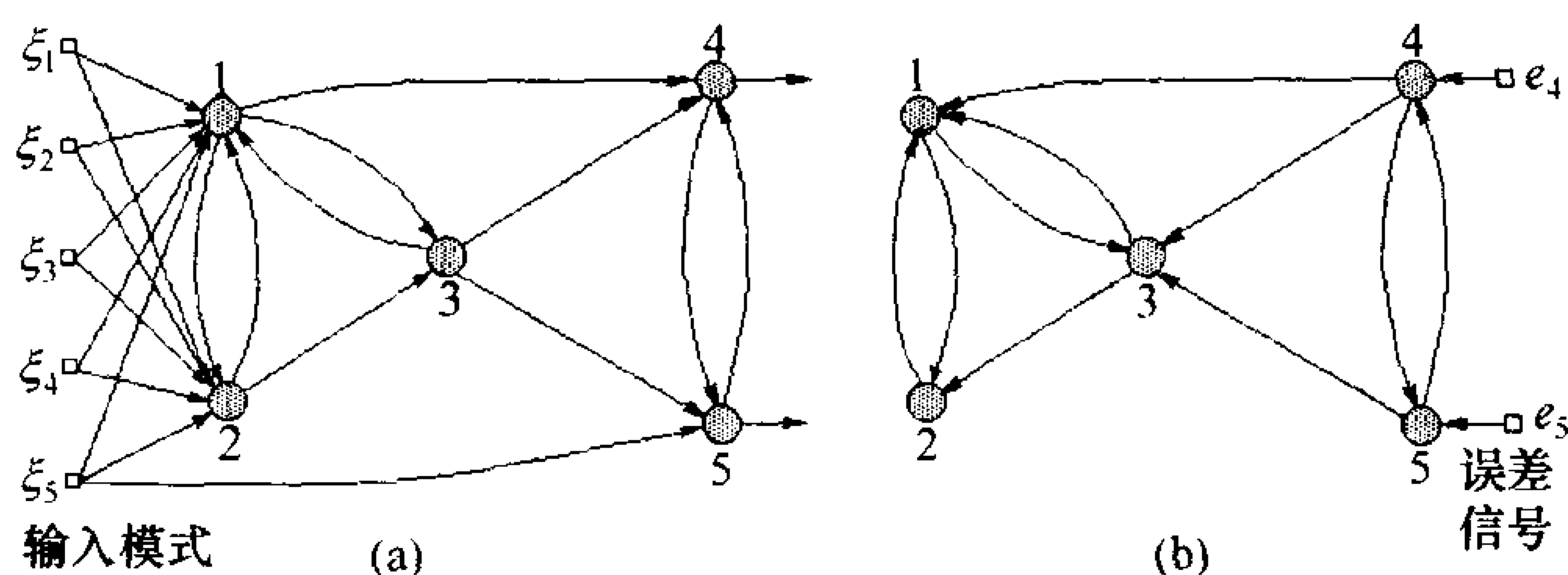


图 3.5.5

网络由输入神经元产生的输出信号稳定的状态值提供了它的反应，假设系统收敛于不动点，那么若神经元 j 是输出的，则反应的结果为稳态值 $x_j(\infty)$，是神经元输出信号. 真正的 $x_j(\infty)$ 与期望的反应作比较，于是有了误差信号 e_j，我们可以给出如下的形式：

$$e_j = \begin{cases} d_j - x_j(\infty), & \text{对输出神经元,} \\ 0, & \text{其它.} \end{cases} \tag{3.5.19}$$

假设输入模式 $\{\xi_j\}$ 与期望模式 $\{d_j\}$ 均为时不变的. 如果神经元 j 是输出，那么方程 (3.5.18), (3.5.19) 是可以应用的，这种情形，称之为 **夹子**. 另一方面，如果神经元 j 是隐层的，则有

$$I_j = e_j = 0, \tag{3.5.20}$$

这种情形，我们称之为 **非夹子** 或 **自由运动**.

我们的目标是寻找一个网络的权矩阵，这个算法是对于最终状态 $x(\infty)$ 和给定模式 $\{d_j\}$，使得吸引子最终的一部分具有期望的 $\{d_j\}$，这可以通过求 ε 的最小值来实现，而成本函数 ε 可以由下式给出：

$$\varepsilon = \frac{1}{2}\sum_k e_k^2, \tag{3.5.21}$$

其中 e_k 由 (3.5.19) 给出，因子 $\dfrac{1}{2}$ 是为了计算方便.

成本函数 ε 通过不动点依赖于网络的权重. 求 ε 最小值动力系统的方法是在权空间中让网络沿学习轨道的演化, 而这一个轨道是沿 ε 的梯度向量的反向下降. 具体地说, 改变调整权重的方法是

$$\Delta w_{rs} = -\eta \frac{\partial \varepsilon}{\partial w_{rs}}, \tag{3.5.22}$$

其中 η 是学习率参数, 它定义了权矩阵改变的时间尺度. 参数 η 必须足够小, 使得权矩阵的变化非常缓慢, 且关于方程的运动行为也是如此, 否则, 成本函数 ε 不再是一个点吸引子的函数.

等式 (3.5.22) 定义了网络在权空间中的演化, 坐标由网络中的权重定义. 随着它的演化, 网络也在相空间中演化. 考虑的网络可以由神经动力学方程

$$\frac{\mathrm{d}v_j(t)}{\mathrm{d}t} = -v_j(t) + \sum_i w_{ji}\phi(v_i(t)) + I_j \quad (j = 1, 2, \cdots, N) \tag{3.5.23}$$

或相关的方程

$$\frac{\mathrm{d}x_j(t)}{\mathrm{d}t} = -x_j(t) + \phi\left(\sum_i w_{ji}x_i(t)\right) + K_j \quad (j = 1, 2, \cdots, N) \tag{3.5.24}$$

定义. 方程 (3.5.23) 可以方便地表示为

$$\frac{\mathrm{d}v_j(t)}{\mathrm{d}t} = -v_j(t) + \sum_i w_{ji}x_i(t) + I_j \quad (j = 1, 2, \cdots, N), \tag{3.5.25}$$

其中 $x_j = \phi(v_j)$. 状态向量 x 的分量构成了相空间的坐标.

等式 (3.5.22) 右边与偏导数 $\dfrac{\partial \varepsilon}{\partial w_{rs}}$ 有关, 由 (3.5.19), (3.5.21) 有

$$\frac{\partial \varepsilon}{\partial w_{rs}} = -\sum_k e_k \frac{\partial x_k(\infty)}{\partial w_{rs}}, \tag{3.5.26}$$

在 (3.5.22) 式中用 (3.5.26) 式, 可得

$$\Delta w_{rs} = \eta \sum_k e_k \frac{\partial x_k(\infty)}{\partial w_{rs}}. \tag{3.5.27}$$

为了得到偏导数 $\dfrac{\partial x_k(\infty)}{\partial w_{rs}}$, 我们采用以下过程 (Pineda, 1988):

(1) 以 $x(\infty)$ 表示环路网络的一个不动点，第 j 个分量 $x_j(\infty)$ 定义为

$$x_j(\infty) = \phi(v_j(\infty)), \tag{3.5.28}$$

这里 $v_j(\infty)$ 是神经元 j 在 $t=\infty$ 时刻的启动势. 由等式 (3.5.25), 有

$$v_j(\infty) = \sum_i w_{ji} x_i(\infty) + I_j. \tag{3.5.29}$$

(2) 应用链式法则，可以把偏导数 $\dfrac{\partial x_j(\infty)}{\partial w_{rs}}$ 表示为

$$\frac{\partial x_j(\infty)}{\partial w_{rs}} = \frac{\partial x_j(\infty)}{\partial v_j(\infty)} \frac{\partial v_j(\infty)}{\partial w_{rs}}. \tag{3.5.30}$$

(3) 由等式 (3.5.28), 有

$$\frac{\partial x_j(\infty)}{\partial v_j(\infty)} = \phi'(v_j(\infty)), \tag{3.5.31}$$

这里 $\phi'(\cdot)$ 是非线性启动函数 $\phi(\cdot)$ 相应于它的自变量的导数.

(4) 由等式 (3.5.29), 得到

$$\frac{\partial v_j(\infty)}{\partial w_{rs}} = \sum_i \left(\frac{\partial w_{ji}}{\partial w_{rs}} x_i(\infty) + w_{ji} \frac{\partial x_i(\infty)}{\partial w_{rs}} \right). \tag{3.5.32}$$

因为网络的权是彼此独立的，可知偏导数 $\dfrac{\partial w_{ji}}{\partial w_{rs}}$ 等于 1, 当且仅当 $j=r$, $i=s$, 或者等于 0, 即

$$\frac{\partial w_{ji}}{\partial w_{rs}} = \delta_{jr}\delta_{is}, \tag{3.5.33}$$

这里 δ_{jr}, δ_{is} 是 Kronecker 记号. 将等式 (3.5.33) 代入 (3.5.32), 关于 i 求和，得到

$$\frac{\partial v_j(\infty)}{\partial w_{rs}} = \delta_{jr} x_s(\infty) + \sum_i w_{ji} \frac{\partial x_i(\infty)}{\partial w_{rs}}. \tag{3.5.34}$$

(5) 偏导数 $\dfrac{\partial x_j(\infty)}{\partial w_{rs}}$ 可以等价地表示为

$$\frac{\partial x_j(\infty)}{\partial w_{rs}} = \sum_i \delta_{ji} \frac{\partial x_i(\infty)}{\partial w_{rs}}, \tag{3.5.35}$$

这里的 δ_{ji} 是 Kronecker 记号.

(6) 对等式 (3.5.30) 的左边应用等式 (3.5.35), 在 (3.5.30) 右端代入 (3.5.31) 与 (3.5.32) 后可以得到

$$\sum_i \delta_{ji} \frac{\partial x_i(\infty)}{\partial w_{rs}} = \phi'(v_j(\infty)) \left[\delta_{jr} x_s(\infty) + \sum_i w_{ji} \frac{\partial x_i(\infty)}{\partial w_{rs}}\right]. \tag{3.5.36}$$

下一步，把 (3.5.36) 中的所有导数组合起来，得到

$$\sum_i L_{ji} \frac{\partial x_i(\infty)}{\partial w_{rs}} = \delta_{jr} \phi'(v_j(\infty)) x_s(\infty), \tag{3.5.37}$$

这里 L_{ji} 定义为

$$L_{ji} = \delta_{ji} - \phi'(v_j(\infty)) w_{ji}. \tag{3.5.38}$$

(7) (3.5.37) 式对 $j = 1, 2, \cdots, N$ 都有定义，因此可以把它用矩阵形式写成线性方程组，有

$$L \frac{\partial x(\infty)}{\partial w_{rs}} = \begin{pmatrix} \delta_{1r} \phi'(v_1(\infty)) \\ \delta_{2r} \phi'(v_2(\infty)) \\ \vdots \\ \delta_{Nr} \phi'(v_N(\infty)) \end{pmatrix} x_s(\infty), \tag{3.5.39}$$

这时 L 是 $N \times N$ 矩阵，第 (j, i) 个分量由 (3.5.38) 式给出. 最后方程 (3.5.39) 两边乘以逆阵 L^{-1}, 保留方程组的第 k 行，我们得到

$$\frac{\partial x_k(\infty)}{\partial w_{rs}} = (L^{-1})_{kr} \phi'(v_r(\infty)) x_s(\infty), \tag{3.5.40}$$

这里 $(L^{-1})_{kr}$ 为逆阵 L^{-1} 的 (k,r) 分量. 等式 (3.5.40) 就是所要求的结果.

我们现在可以把 (3.5.40) 代入 (3.5.27) 后, 计算改变量 Δw_{rs}, 结果由 δ 律的简单形式给出, 即

$$\Delta w_{rs} = \eta\delta_r(\infty)x_s(\infty), \tag{3.5.41}$$

这里 $\delta_r(\infty)$ 是局部梯度, 定义为

$$\delta_r(\infty) = \phi'(v_r(\infty))\sum_k e_k(L^{-1})_{kr}. \tag{3.5.42}$$

注意, 切勿将此 δ 与 Kronecker 记号混淆. 等式 (3.5.41), (3.5.42) 形式上具体给出了修改网络权重的一个新的学习律. 不幸的是, 为了计算 $\delta_r(\infty)$, 还需要计算矩阵 L 的逆阵, 而矩阵直接求逆要牵涉到整体的计算, 所以当前形式的学习算法对于应用到神经网络中是不合适的.

我们可以用一种 "局部" 方法来计算 $\delta_r(\infty)$, 通过引入一个相关的动力系统, 以克服这个限制. 为了能够这样做, 定义

$$\delta_r(\infty) = \phi'(v_r(\infty))y_r(\infty), \tag{3.5.43}$$

这里

$$y_r(\infty) = \sum_k e_k(L^{-1})_{kr}. \tag{3.5.44}$$

接下来并不求 (3.5.44) 中的逆阵, 而是对新的变量 $y_r(\infty)$ 提取出一组线性组合

$$\sum_r L_{rj}y_r(\infty) = e_j. \tag{3.5.45}$$

然后应用由 (3.5.38) 给出的 L_{ij} 的定义, 将 (3.5.45) 写为

$$\sum_r \left[\delta_{rj} - \phi'(v_r(\infty))w_{rj}\right] y_r(\infty) = e_j,$$

由 δ_{rj} 的定义, 可得到

$$y_j(\infty) - \sum_r \phi'(v_r(\infty))w_{rj}y_r(\infty) = e_j. \tag{3.5.46}$$

为了使表达式比较清楚，我们对等式 (3.5.46) 做两件事： (i) 把指标 r 代为 i, (ii) 把项重排表示为

$$0 = -y_j(\infty) + \sum_i \phi'(v_i(\infty))w_{ij}y_i(\infty) + e_j \quad (j = 1, 2, \cdots, N), \tag{3.5.47}$$

这里为了完备化，指标 j 的范围已经增大了.

最后，得到线性方程组 (3.5.47) 的解实际上是相关动力系统

$$\frac{\mathrm{d}y_j(t)}{\mathrm{d}t} = -y_j(t) + \sum_i \phi'(v_i(\infty))w_{ij}y_i(t) + e_j \tag{3.5.48}$$

的不动点 $y_i(\infty)$ 的计算，因此也是 $\delta_j(\infty)$ 的计算，这里已经与矩阵求逆无关了. 注意到，从 (3.5.47) 到 (3.5.48) 这一步并不唯一，(3.5.47) 可以以不同的方式到达相关的微分方程.

一种特殊情况是 $\phi(\cdot)$ 由 Logistic 函数定义， $\phi'(\cdot)$ 的计算可以用关系

$$\phi'(v_i(\infty)) = x_i(\infty)[1 - x_i^2(\infty)], \tag{3.5.49}$$

这里的 $x_i(\infty)$ 是不动点 $x(\infty)$ 的第 i 个分量，由神经元 i 的稳态输出表示，并由 (3.5.48) 的导出完成了新的学习算法. 为了有一个物理理解，我们得到以下重要的观察结果：

(1) 由 (3.5.25) 可以看到给定的回路网络由权重集合 $\{w_{ji}\}$ 所决定，由计算状态变量 $\{x_j\}$ 的集合来构造网络结构;

(2) 由 (3.5.48) 可以看到计算 $\{y_i(t)\}$ 的网络可以通过重新安排给定的网络来得到：

1) 神经元 i 到神经元 j 的权重被 $\phi'(v_i(\infty))w_{ij}$ 所替代. 换句话说，通过网络的信号传递是反向的，得到的网络与原始网络相伴随.

2) 原始网络的输出神经元假设为伴随网络的输入神经元. 这些“新”的输入神经元的偏差项被设为 e_j, e_j 是一个由 (3.5.21) 定义的误差信号.

3.5.4 离散网络通过权矩阵的选取引出的复杂动力学

这一节仍回到前面已经讨论过的简单的离散网络，它的数学模型为

$$x_{it} = f\left(\sum_{j=1}^{N} w_{ij} x_{j,t-1}\right), \quad i = 1, 2, \cdots, N, \tag{3.5.50}$$

其中 x_{it} 为第 i 个神经元 (或称为节点) 在时刻 t 取实值的输出，w_{ij} 为从节点 j 到节点 i 的实值权重值，f 是 S 型函数，例如

$$f(u) = \frac{1}{1 + \mathrm{e}^{-u}}. \tag{3.5.51}$$

前面对于吸引子的动力学性质已经显示了该离散网络的收敛及振荡的动力学行为. 那时对权重的选取只是从它是对称的或者对角元为非负的角度来讨论的，很自然会想到实际上生物神经网络系统权重矩阵是非对称的，是有很大的选择自由度的，那么在以很随便的方式来选择权重时，整个网络的吸引子将具有复杂的动力学行为，例如会出现混沌奇异吸引子.

由系统 (3.5.50) 所产生的瞬间模式的广阔的视野图 (如图 3.5.6 所示)，表示了吸引子的能量谱的复杂度的类似熵的变量. 吸引子在由 20 个节点组成的网络及随机的 225 个参数权空间的两个参数的子空间生成的 273 (21×13) 个点上显示了自己的形态. 两个参数是下式中的 r 和 a:

$$w_{ij} = r\sigma(a)S_{ij} + (1 - \sigma(a))A_{ij}, \tag{3.5.52}$$

式中 S_{ij} 是对称矩阵 S 的元素，A_{ij} 是反对称矩阵 A 的元素，$\sigma(a)$ 是 a 的函数，随 a 从 -1 到 1 变化，$\sigma(a)$ 从 0 变到 1，当 a 接近 -1 或 1 时变化最快，因而当 $a = -1$ 时，W 是纯粹的反对称阵，当 $a = 1$ 时，W 是纯粹的对称阵，当 $a = 0$ 时是随机的. (3.5.51) 在 0 的斜率的倾斜度随参数 r 增长，当 $r \to 0$ 时为线性的网络，当 $r \to \infty$ 时为高阈值情形，平均值用 “•” 表示，方差用 “⊙” 的环表示，这两个值通过权矩阵在节点的 10 个初始状态的随

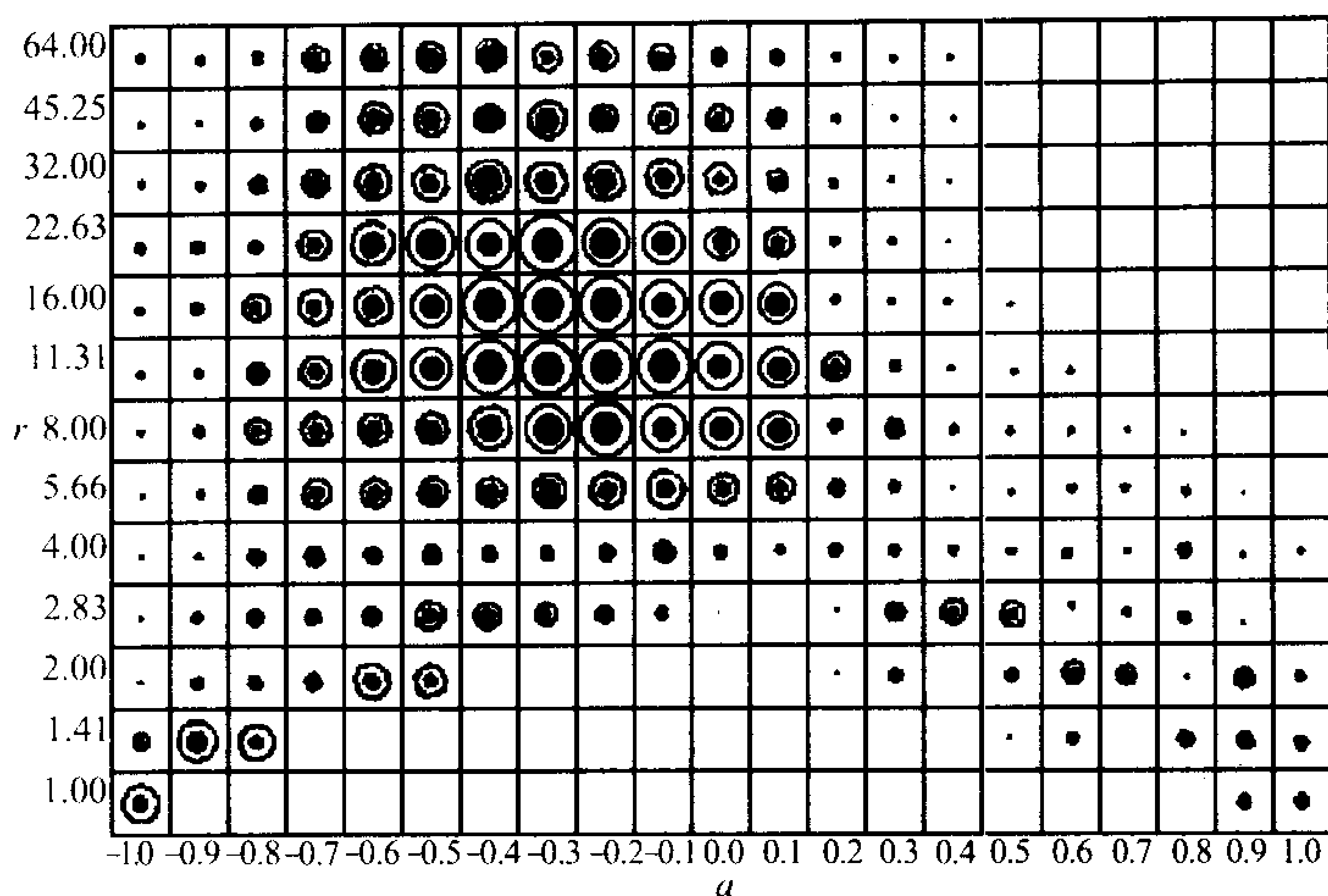


图 3.5.6　权重的对称性与数量对动力学行为复杂性的影响图

20 个节点的网络, $\Delta t = 1$, 最大的能量均值为 -33.887db, 最大的能量方差为 -36.318db, 最大的平均熵为 3.774, 最大的熵的方差为 2.311

机值而得到. 我们可以看到最复杂的行为集中于 r 取某些特定的范围时的轻度反对称阵矩阵. 较小的 r (几乎是线性网络) 所对应的空白面积代表不动点行为, 在 $a = 1$ 处的空白 (对称矩阵) 相应于不动点和 2 周期振荡 (没有被使用的变量所识别), Frumkin 和 Moses (1986, Physical Review) 证明了这是当 $r \to \infty$ 时纯对称权阵唯一可能的运动. 纯反对称的情形 ($a = -1$) 表示不动点和周期 4 极限环, $r \to \infty$ 时, 反对称情形有 Gutfreund, Reger 和 Young (1988, J. of Physics A.) 的相应定理作依据. 他们对于 4 至 25 个节点的网络得到了类似的定性的结论, 随节点的数目的增长而导致的整个复杂度的增长 (粗糙的线性), 与随着最大的复杂度的区域的增加而导致的增长是一样的 (为实现这一稳定性, 随机权重的量值要与节点数目的平方根作相反的改变, 从而保持在随机状态下到每个节点中的 RMS 期望的输入).

以下通过研究离散系统与连续系统的联系, 来比较复杂一些的离散网络 (3.5.50) 的动力学行为. 考虑连续的网络

$$\frac{\mathrm{d}y_i(t)}{\mathrm{d}t} = -y_i(t) + f\left(r\sum_j w_{ij}y_j(t)\right), \quad i = 1, 2, \cdots, N, \quad (3.5.53)$$

相应的离散网络为

$$y_i(t+\Delta t) = (1-\Delta t)y_i(t) + \Delta t f\left(r\sum_j w_{ij}y_j(t)\right). \qquad (3.5.54)$$

上式当 $\Delta t = 1$ 时, 则变为 (3.5.50); 当 $\Delta t \to 0$ 时, 则变为 (3.5.53).

Renals (1990, Vol.412 of Lecture Notes in Computer Science) 得出当 Δt 减少时, 最大复杂度逐渐增加, 最大复杂度的区域向更高的 r 值和更低的 a 值 (更加反对称) 的方向移动. 不动点演化所覆盖的区域, 是受项 $(1-\Delta t)y_i(t)$ 的演化的影响. 他又报告了连续系统 (3.5.53) 有较低的 r 处的不动点的行为. Kiiten 和 Clark (1986, Physics Letters) 还发现了类似于 (3.5.53) 的一个系统中的混沌行为, 但要求到每一个节点的连接是有限的. 在较高的 r 和较小的 Δt (< 0.1) 所对应的区域的研究不是很理想, 因为随着 Δt 的较小复杂度的光滑变化还没有令人信服的事实, 因此我们还不能确定它的连续极限是有意义的.

Renals (1990, J. of Statistical Physics) 也研究了利用分叉图吸引子剖面、分维数计算、 Lyapunov 指数来考察的一些复杂系统的吸引子的结构问题.

图 3.5.7 表示了吸引子如何随 r 的变化而发生改变. 对 r 轴上的每一个点, 从同一个初始点直到过渡时间结束, 一个随机的 8 节点的网络迭代 10000 个时间步长, 在每一步迭代之后, 每个特定的节点的输出画在 y 方向上, 无论选取哪一个节点考虑方程 (3.5.50) 与 (3.5.53) 都很容易理解. 这个不动点随 r 的增加而连续地变化, 直到 $r = 5.2822$ 时出现复杂的运动行为. 这里系统的状态点以大约是映射 (3.5.50) 的每次迭代的 0.26 个转数的有理频率沿拓扑上圆的路径运动, 这个区域中分支图的放大视图见图 3.5.8.

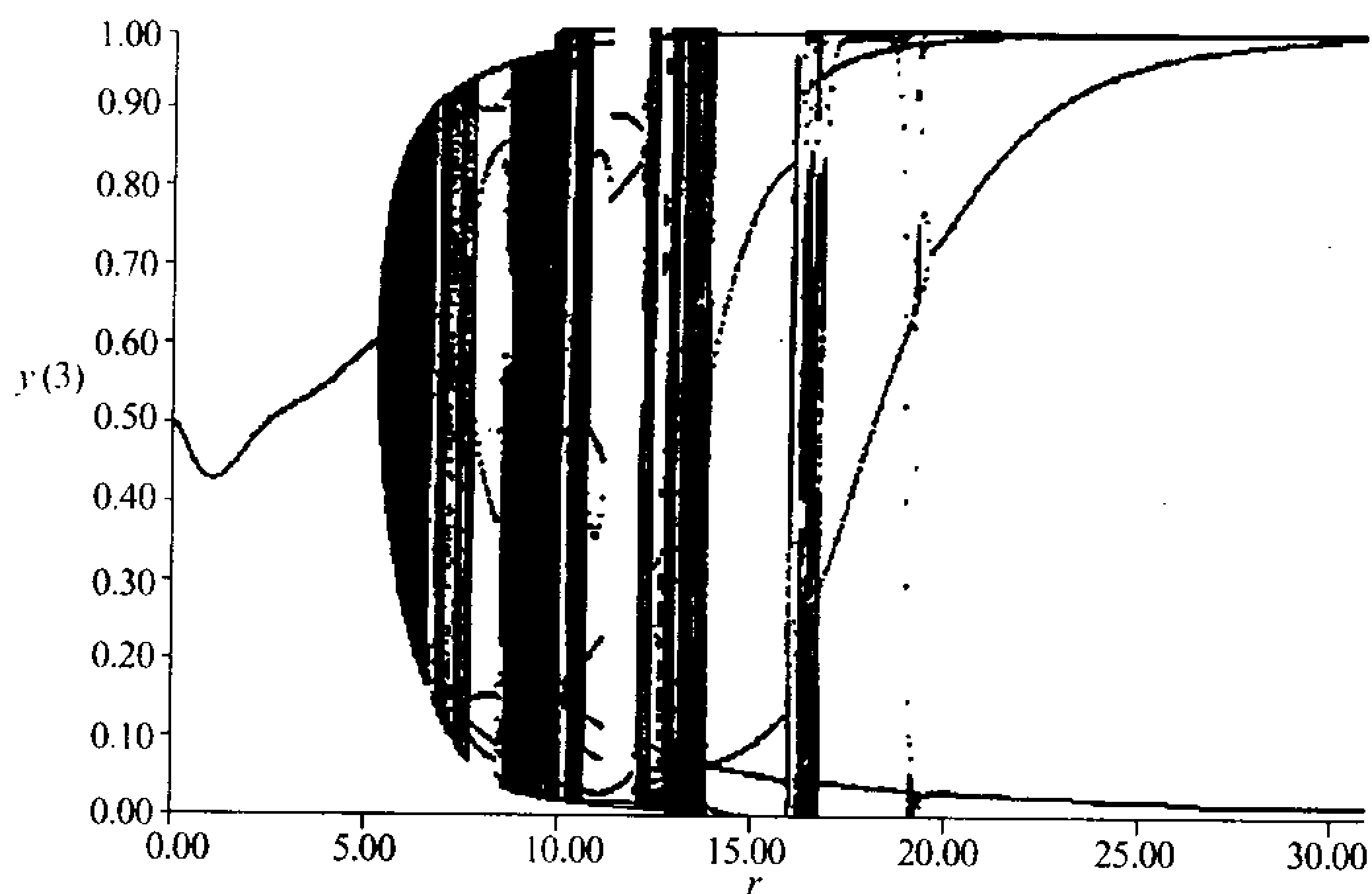
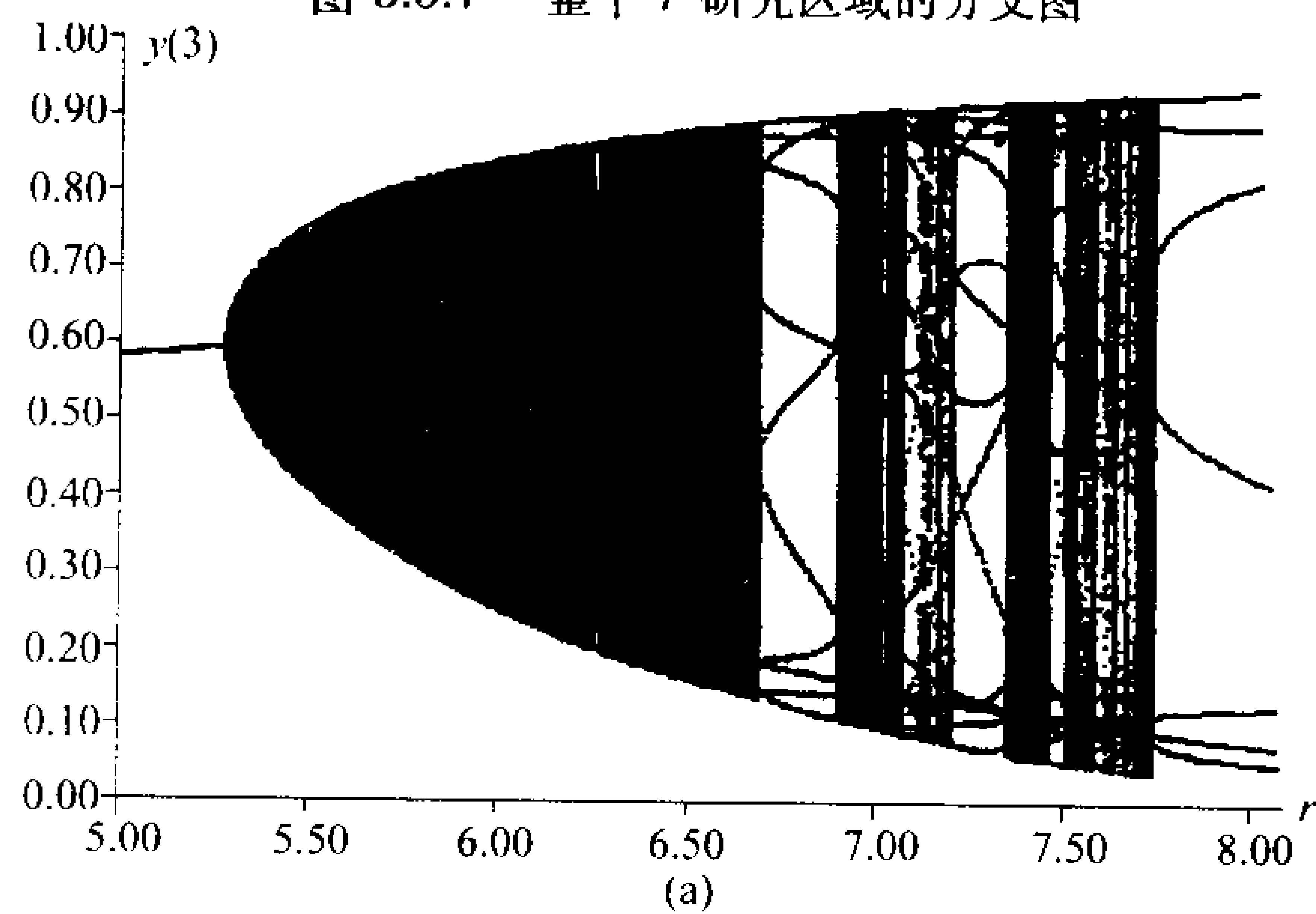


图 3.5.7　整个 r 研究区域的分支图



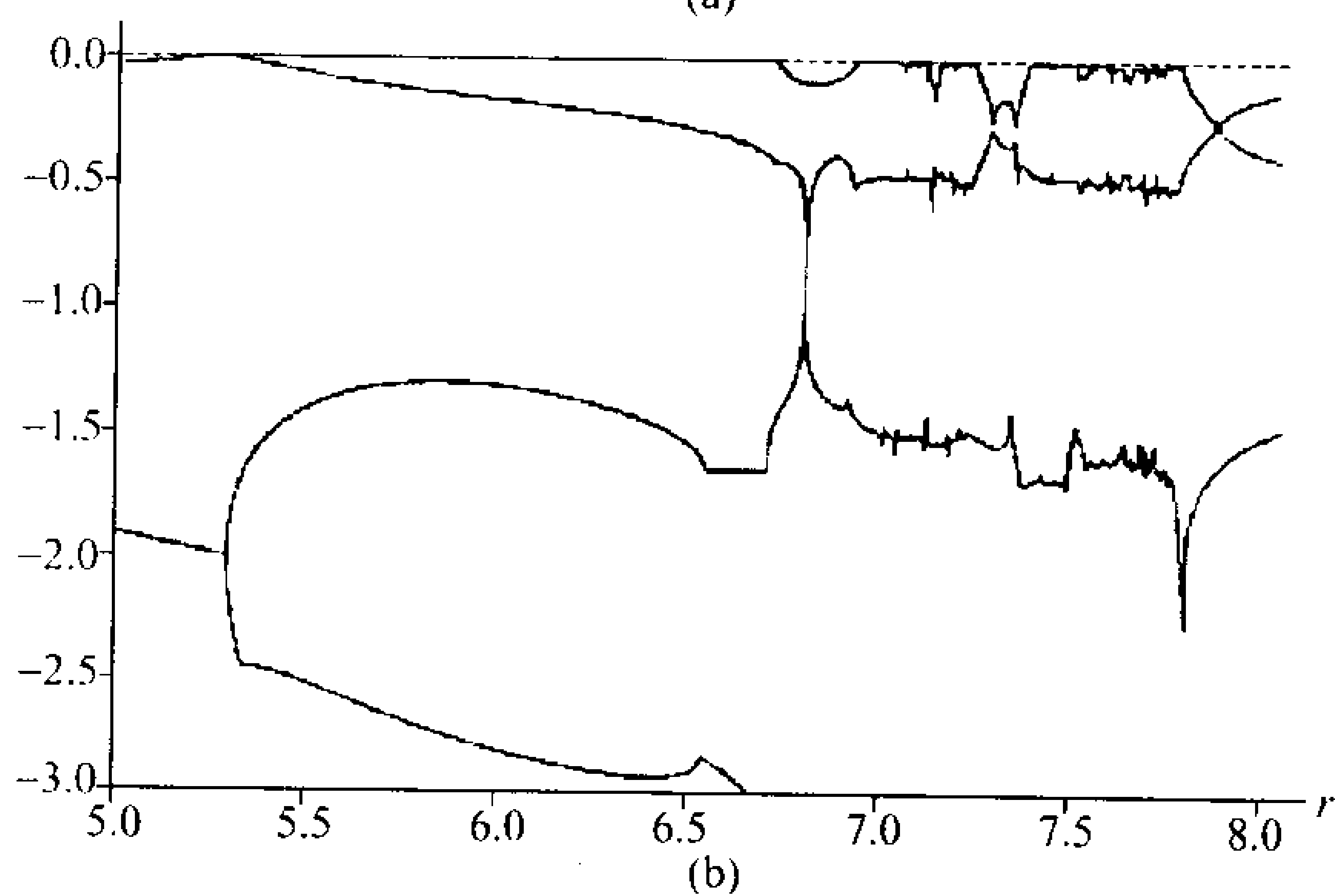


图 3.5.8　(a) 随 r 变化的 $y(3)$ 的倍周期分支图, (b) 相应的 Lyapunov 指数

另外一个成比例频率出现在 $r = 5.2840$ 处，在分叉后迅速呈现振荡行为，这两个频率连续变化，当它们的频率彼此之间以及与映射的周期之间成比例时，锁模形成极限环．还有一种频率大约在 $r = 6.9$ 处带有一个更大范围的谱而出现，吸引子的部分结构开始出现反折叠以及更混沌的行为，这种模式与 Ruelle-Takens-Newhouse 通向混沌的道路是相符合的．在 $r = 10.69$ 处，有一个

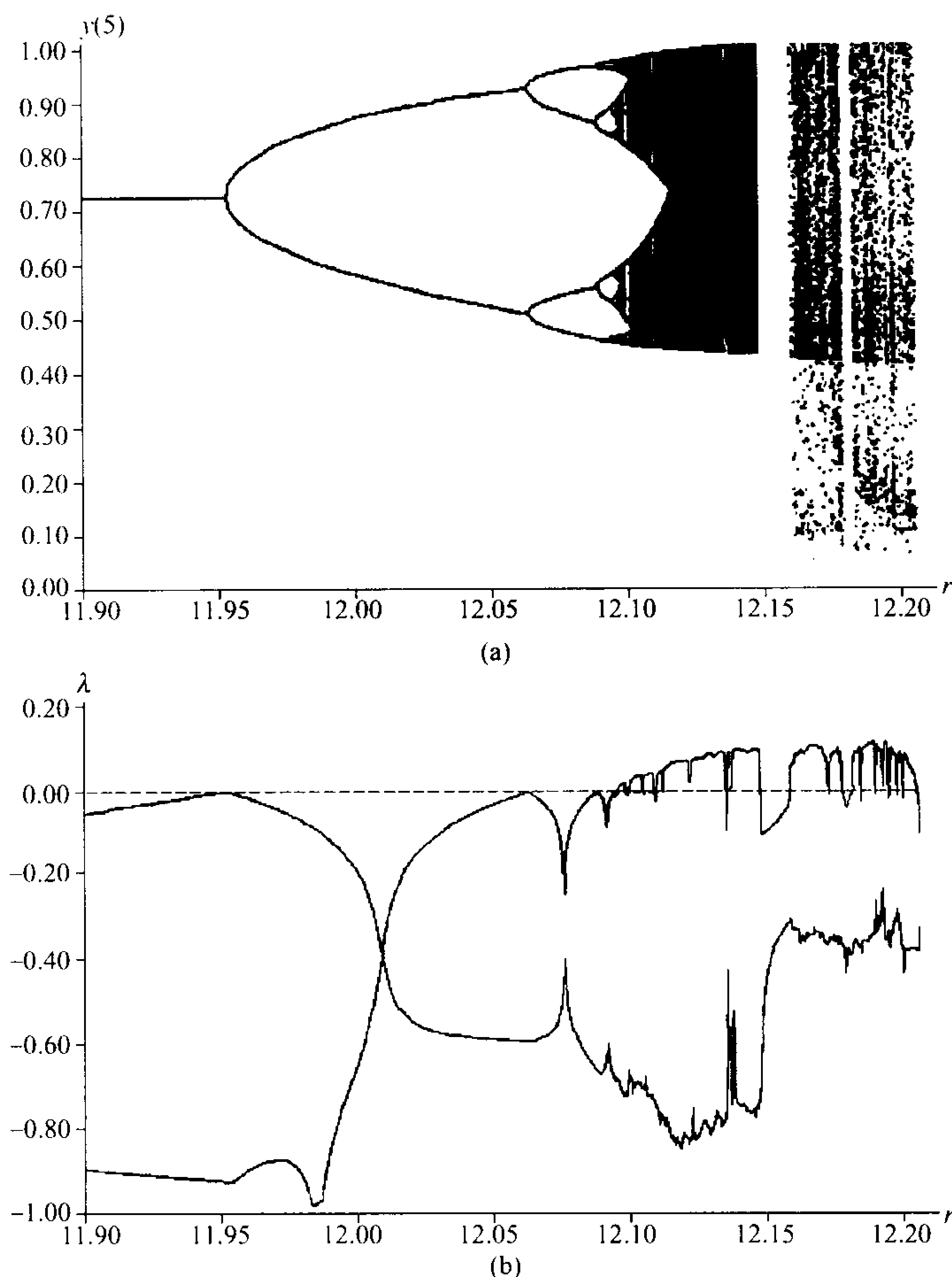


图 3.5.9　(a) 随 r 变化的 $y(5)$ 的倍周期分支图，(b) 相应的 Lyapunov 指数

相关维近似为 2.2 的混沌吸引子, 对于不动点和极限环, 最大 Lyapunov 指数为负的, 当没有锁模时, 增长为 0.

周期倍分叉通向混沌的道路可以在图 3.5.9 中 $r = 12$ 处明显地看出, 正确的普适的 Feigenbaum 常数从这个图可以计算出来.

§3.6 由 H–H 方程及 Chay 方程描述的神经活动中的各类动力学行为

神经中枢的单个神经元的细胞内外记录主要偏重于大神经元, 该神经元一般具有轴突性, 并能产生作用电压, 因此神经系统中单个神经元的活动可以看成一个系列的完全作用或非完全作用电压 [20]. 在麻醉制备品中, 神经元是静止的, 并且作用电压必须用合适的输入来激发, 神经元看作临界装置. 非麻醉制备品的单个神经元的长期记录表明大多数神经元表现为一个有背景的不断放电的过程, 该过程可以通过输入得以改善. 神经元膜的生物物理实验已鉴别出产生作用电压的过程, 并发现神经膜的电学行为可以用非线性微分方程来描述. 这里讨论由这些兴奋方程所表现的动力学行为以及这些行为间的分叉, 所考察的兴奋方程为

(1) Hodgkin–Hoxley 方程 [18] (H–H 方程), 它是描述鱿鱼巨轴突细胞膜的性质的方程;

(2) Chay 方程 [19], 它是描述神经体膜的生物物理兴奋方程的简化形式.

3.6.1 Hodgkin–Huxley 方程

H–H 方程于 1952 年提出, 它给出了描述鱿鱼巨轴突细胞膜表面电流与电压时间的依赖关系. 当电压不随距离发生变化时, 通过电容的电流与离子电流数值相等但符号相反. $C_m \dfrac{\mathrm{d}V}{\mathrm{d}t} = I_i$, 这里离子电流 I_i 是钠离子电流 I_{Na}, 钾离子电流 I_{K} 和漏电 I_{L} 之和. 每个电流是电导率 $g_{\mathrm{Na}}, g_{\mathrm{K}}, g_{\mathrm{L}}$ 和电化学梯度 $(V - V_{\mathrm{Na}}), (V - V_{\mathrm{K}})$,

$(V-V_{\mathrm{L}})$ 的乘积. 漏导 g_{L} 是一个常数, 而 g_{Na} 和 g_{K} 随电压 V 和时间而变化.

$$\begin{aligned} I_i &= I_{\mathrm{Na}} + I_{\mathrm{K}} + I_{\mathrm{L}} = F(V,m,n,h) \\ &= g_{\mathrm{Na}}(V-V_{\mathrm{Na}}) + g_{\mathrm{K}}(V-V_{\mathrm{K}}) + g_{\mathrm{L}}(V-V_{\mathrm{L}}) \\ &= m^3 h \overline{g}_{\mathrm{Na}}(V-V_{\mathrm{Na}}) + n^4 \overline{g}_{\mathrm{K}}(V-V_{\mathrm{K}}) + g_{\mathrm{L}}(V-V_{\mathrm{L}}), \end{aligned} \tag{3.6.1}$$

这里 $\overline{g}_{\mathrm{Na}}, \overline{g}_{\mathrm{K}}$ 是钠、钾的最大电导率, $0<m,n,h<1$, m 和 n 是激活变量, h 是非激活变量, 它们都满足下列一阶常微分方程组

$$\begin{cases} \dfrac{\mathrm{d}m}{\mathrm{d}t} = \alpha_m(1-m) - \beta_m m, \\ \dfrac{\mathrm{d}n}{\mathrm{d}t} = \alpha_n(1-n) - \beta_n n, \\ \dfrac{\mathrm{d}h}{\mathrm{d}t} = \alpha_h(1-h) - \beta_h h, \end{cases} \tag{3.6.2}$$

其中系数 α_m, β_m, α_n, β_n, α_h, β_h 是电压的函数. 该系统是用来描述当电压固定在维持值时的电流. 这些方程的解在一个较宽的条件下可产生鱿鱼巨轴突的电学行为. H–H 方程不是一个通用方程, 它表示一种特殊膜的行为, 而其它的神经和肌肉膜可用一个不同的特殊兴奋方程描述 (Holden, 1981).

3.6.2 Chay 方程

不同神经和肌肉细胞的兴奋膜的电压箝位分析显示一个宽带的传导过程, 都有 H–H 型动力学特征, 并且 α 和 β 依赖于电压. 另外它们都有依赖于电导率的电荷, 以这种电导率电荷 Ca^{++} 的进入 (时间电流) 使钾离子 K^{+} 选择的电导率起作用. Chay 将 5 个变量的 β 细胞模型简化, 而 Chay 和 Rinzel 给出了 3 个变量的

软体神经体膜的模型

$$
\begin{cases}
\dfrac{\mathrm{d}V}{\mathrm{d}t} = g_1^* m_\infty^3 h_\infty (V_1 - V) + g_{\mathrm{K},V} n^4 (V_{\mathrm{K}} - V) \\[2ex]
\qquad\qquad + g_{\mathrm{K},C}^* \dfrac{C}{1+C}(V_{\mathrm{K}} - V) + g_{\mathrm{L}}^*(V_{\mathrm{L}} - V) + I, \\[2ex]
\dfrac{\mathrm{d}C}{\mathrm{d}t} = \rho\{m_\infty^3 h_\infty (V_C - V) - k_C C\}, \\[2ex]
\dfrac{\mathrm{d}n}{\mathrm{d}t} = \dfrac{n_\infty - n}{\tau_n}.
\end{cases}
\tag{3.6.3}
$$

时间 t (以微秒计) 是独立变量，非独立变量有膜电压 V (以 mV 计), 无量纲钙的浓度 C, 激活概率 n.

如果令 y 代表 h, m 或 n, 则我们可明确表示 h_∞, m_∞ 和 n_∞,

$$
y = \frac{\alpha_y}{\alpha_y + \beta - y},
$$

其中

$$
\begin{aligned}
\alpha_h &= 0.07\exp(9 - 0.05V - 2.5), \\
\beta_h &= \frac{1}{1+\exp(-0.1V - 2)}, \\
\alpha_m &= \frac{0.1(25+V)}{1-\exp(0.1V - 2.5)}, \\
\beta_m &= 4.0\exp\left(-\frac{V+50}{18}\right), \\
\alpha_n &= \frac{0.01(20+V)}{1-\exp(-0.1V - 2.0)}, \\
\beta_n &= 0.125\exp\left(-\frac{V+30}{80}\right),
\end{aligned}
\tag{3.6.4}
$$

另外 $\tau_n = \dfrac{1}{230(\alpha_n + \beta_n)}$.

有了这些定义我们得到

$$\frac{\mathrm{d}n}{\mathrm{d}t} = 230((1-n)\alpha_n - n\beta_n),$$

其参数为用 mV 计量的电流 I 和钙 Ca^{++} 对 K^+ 敏感通道的最大电导率 g^*_{KC}. 该通道被薄膜电容所分割，其值为

$$g^*_1 = 1800\mathrm{s}^{-1}, \quad g^*_{\mathrm{K},V} = 1700\mathrm{s}^{-1}, \quad g^*_{\mathrm{L}} = 7\mathrm{s}^{-1},$$
$$V_1 = 100\mathrm{mV}, \quad V_{\mathrm{K}} = -75\mathrm{mV}, \quad V_{\mathrm{L}} = 40\mathrm{mV}, \quad V_C = 100\mathrm{mV},$$
$$k_C = 3.3/18\mathrm{mV}, \quad \rho = 0.27\mathrm{mV}^{-1}\ \mathrm{s}^{-1}.$$

3.6.3 平衡点的分枝

分枝点定义为在参数控制空间中结构非稳向量场中的一点. 我们讨论 3 种分枝点，即鞍点、 Hopf 分枝和倍周期点.

考察非线性微分方程

$$\ddot{x} + \dot{x} + x^2 - \mu = 0,$$

它可写成两个一阶的微分方程系统

$$\begin{cases} \dot{x} = y, \\ \dot{y} = \mu - x^2 - y \end{cases}$$

或两个非耦合的微分方程形式

$$\begin{cases} \dot{x} = \mu - x^2, \\ \dot{y} = -y, \end{cases}$$

这里 x 是标量， x 的奇异点可以由 $\dot{x}=0, \dot{y}=0$ 解得. 奇异点在空间 (x,y,μ) 中的轨迹是一个平面 $y=0$ 的抛物线. 对于 $\mu<0$, x 没有奇异点，而对于 $\mu>0$, x 有两个奇异点 $(\pm\sqrt{\mu},0)$, 其 Jacobi

矩阵是

$$J = \begin{pmatrix} -2x & 0 \\ 0 & -1 \end{pmatrix},$$

当 $\mu > 0$ 时，它在 x 的奇异点有非零特征根 (见图 3.6.1). Jacobi 矩阵特征根满足 $\det(J - \lambda I) = 0$, 这里 I 是单位矩阵.

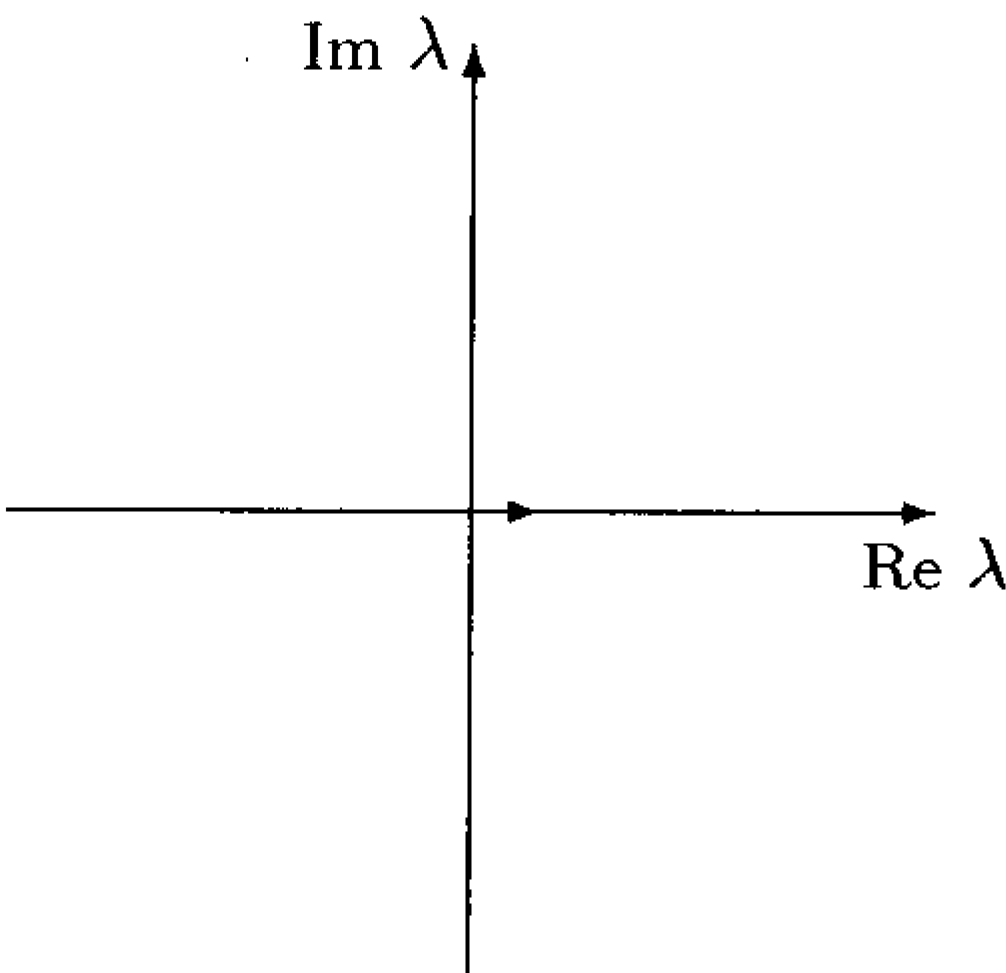


图 3.6.1　复 λ 平面中的实非零特征值

有一个鞍点 $x = -\sqrt{\mu}$ 和一个结点 $x = \sqrt{\mu}$. 随着控制参数 μ 的减少，这两个平衡点彼此接近，且在 $\mu = 0$ 时它们碰撞成一个鞍结点，该点是一个结构稳定的分枝点.

对于 $\mu < 0$, 平衡点消失且每条轨线趋于无穷. 表示一个兴奋膜的微分系统包括一系列的参数，它们可能包括细胞内外介质的离子浓度、最大离子率或者离子通道密度、生化制剂和药物制剂的浓度以及外加电流. 作为一个实例，考察外加电流密度 I 作为参数的方程

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -f(V) - I,$$

这里 $f(V)$ 由 H–H 方程给出.

H–H 方程是 4 维的微分方程组，用分析的办法很难求解，只能用数值积分的方法求得. 为了讨论解的动力学行为，可以通过检

验不动点来进行，即 $\frac{\mathrm{d}V}{\mathrm{d}t}, \frac{\mathrm{d}m}{\mathrm{d}t}, \frac{\mathrm{d}n}{\mathrm{d}t}$ 和 $\frac{\mathrm{d}h}{\mathrm{d}t}$ 全部为零，满足

$$f(V, m_\infty(V), n_\infty(V), h_\infty(V)) + I = 0,$$

这里 $m_\infty(V) = \frac{\alpha_m(V)}{\alpha_m(V) + \beta_m(V)}$, $n_\infty(V)$, $h_\infty(V)$ 同样可得. H–H 方程组的相应 Jacobi 矩阵

$$J = \begin{pmatrix} -\frac{\partial f}{\partial V} & -\frac{\partial f}{\partial m} & -\frac{\partial f}{\partial n} & -\frac{\partial f}{\partial h} \\ \lambda_m m'_\infty & -\lambda_m & 0 & 0 \\ \lambda_m n'_\infty & 0 & -\lambda_n & 0 \\ \lambda_m h'_\infty & 0 & 0 & -\lambda_h \end{pmatrix}$$

的特征值可以在任何不动点用数值方法估计，这里 $\lambda(V) = \alpha(V) + \beta(V)$ 对 m, n 和 h 都成立. 在某些条件下，稳定解点可形成一条曲线，该曲线表明有多个平衡点：对某些特殊的 I 值，满足条件的 V 不唯一. 多个平衡点产生了鞍结分枝点，该点有一个特征值为零，在这一分枝点平衡解的稳定性发生变化. 如果 Jacobi 矩阵在不动点的值都具有负实部，则该不动点是线性稳定的，且平衡点对应一个不动膜.

H–H 方程不动点的稳定性可因影响特征多项式系数的任何一个参数的变化而改变. 由于参数是单调的，特征值在不动点失稳时发生变化，此时某些参数处于临界值. 从一个稳定不动点到非稳定不动点的变化发生在某些特征值的实部与虚轴相交处，在由分枝参数的临界值所确定的分枝点处，分枝点可能发生在不动点的不同曲线间、周期解间或不动点与周期解间.

为了获得 H–H 方程平衡点的分枝点，有必要改变两个参数，如 $\bar{g}_\mathrm{K}$ 和 I 或 $[\mathrm{Ca}^{++}]$ 和 I.

对 $\bar{g}_\mathrm{K} > 10.6\mathrm{ms{\cdot}cm^{-2}}$, 任意 I 仅存在唯一的平衡解；对 $\bar{g}_\mathrm{K} < 10.6\mathrm{ms{\cdot}cm^{-2}}$, 在 $\bar{g}_\mathrm{K}$–I 平面某个区域内会出现多个平衡点，它是分

枝点的投影，在这个区域内可以用数值积分发现 V 在平衡点或在最小平衡点而不是在不稳定的平衡点. 多平衡点区域的边界是平衡点间的分枝.

3.6.4 Hopf 分枝

Hopf 分枝 是从不动点到小振幅周期解，若发生在一对共轭复根穿越虚轴时，则速度是非零的. 在 Hopf 分枝点出现的小振幅解在超分枝点稳定，而在次分枝点不稳定.

考察非强迫阻尼的非线性振子

$$\ddot{x} + \mu\dot{x} + x + f(x, \dot{x}) = 0,$$

这里 $f(x, \dot{x})$ 含有非线性项.

这个二阶方程可以写成两个一阶微分方程

$$\begin{cases} \dot{x} = y, \\ \dot{y} = \mu y - x - f(x, y). \end{cases}$$

在线性化系统中令 $\dot{x} = 0, \dot{y} = 0$ 可得奇异点，很明显奇异点在原点出现，其 Jacobi 矩阵为

$$\begin{pmatrix} 0 & 1 \\ -1 & \mu \end{pmatrix},$$

其特征值很容易计算.

当 μ 由正到负发生变号时，线性化系统的特征根越过虚轴，该虚轴为在 $\lambda = i, \overline{\lambda} = -i$ 的一对直线 (见图 3.6.2). 对于上述系统，当 $\mu < -2$ 时，原点是渐近稳定的结点；当 $\mu = -2$ 时，原点变成一个稳定螺线；当 $\mu = 0$ 时存在分枝点；当 $\mu > 0$ 时，原点变成一个不稳定的螺线，趋于极限环；当 $\mu > 2$ 时原点是一个不稳定的结点.

幅.

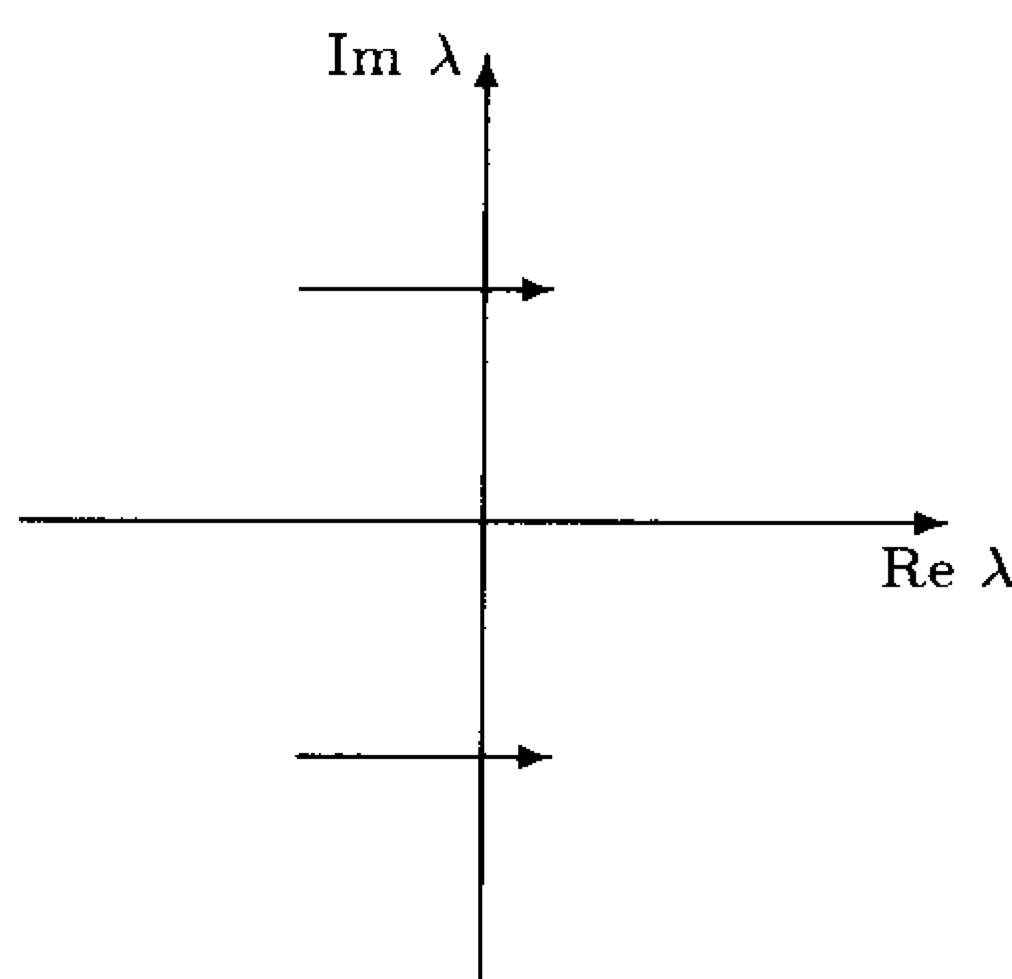


图 3.6.2　复 λ 平面中靠近虚轴的复共轭对的特征值

标准的 H–H 方程 (以电流密度为分枝参数) 的 Hopf 分枝点在文献 [21] 中给出, 在文献 [22,23] 中由数值方法确定. 图 3.6.3 中的 I_1, I_2 接近于 9.78, 154.53μA·cm^{-2}.

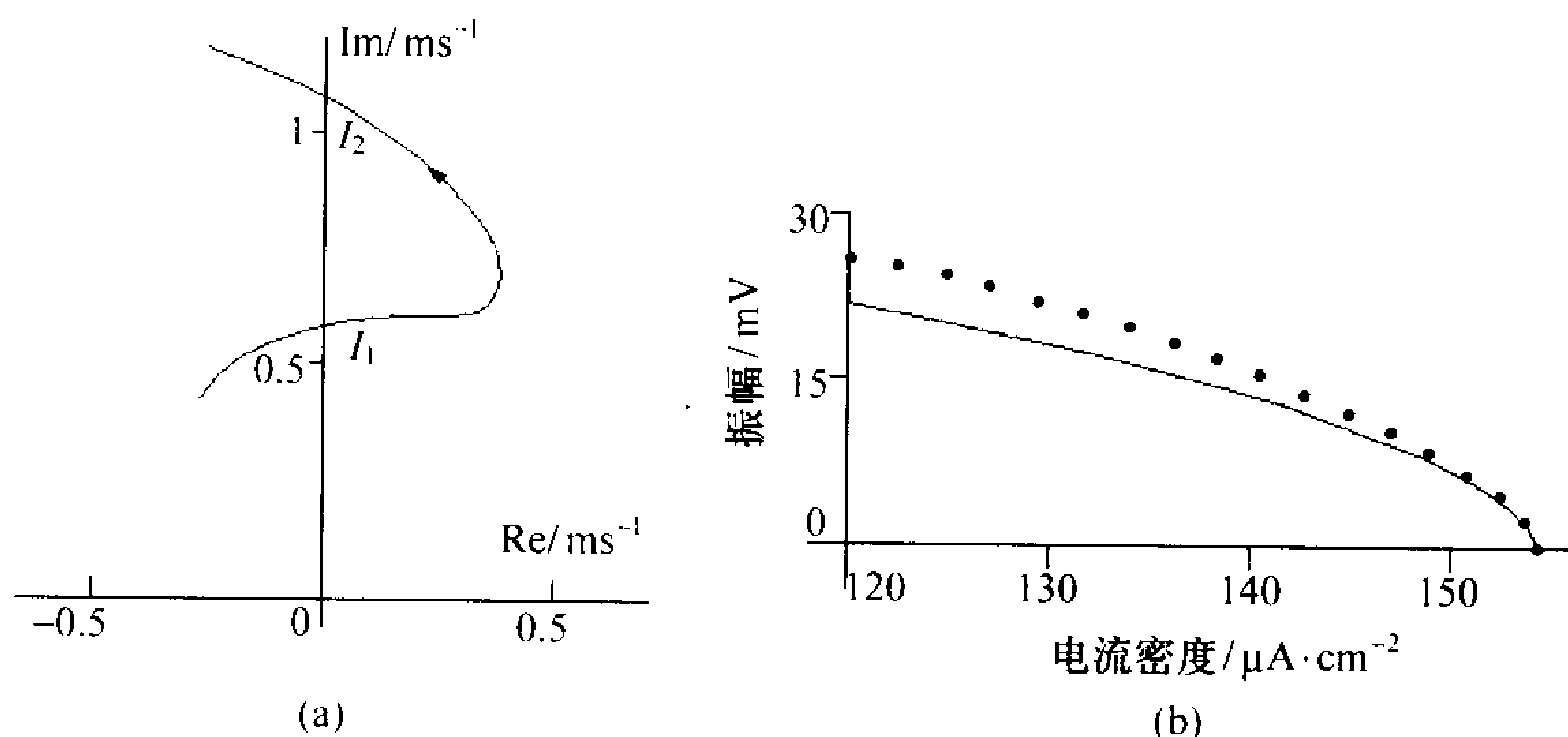


图 3.6.3

(a) H–H 方程一对共轭复特征值之一的轨迹, 图中 I_1, I_2 是 Hopf 分枝点; (b) 在 I_2 附近的周期解的振幅, 图中点由数值积分得到, 线由 Hopf 分支理论得到

在这样的 Hopf 分枝点附近, 小振幅周期解的周期和振幅可以估计出来, 这些值和数值积分所得的结果仅当接近分枝点时才符合得很好. 图 3.6.3(b) 表明在电流密度极接近 I_2 时的周期解的振幅.

对 $I < I_1$ 时, 平衡解是稳定的, 对 $I_1 < I < I_2$ 时, 是不稳定的, 而对 $I > I_2$ 时, 又稳定. 在两个 Hopf 分枝点之间, 唯一的稳定点是周期的. 在 I_2 附近与作用电势的不断放电相对应存在一个

大幅度解；随 I 的增加，该解的振幅不断减少到 I 处消失，如图 3.6.3(b) 所示.

然而，对 I 的值恰比 I_1 小时，$6.25 < I < I_1 = 9.78\mu\text{A}\cdot\text{cm}^{-2}$，平衡解与大振幅周期解都是稳定的，而这些稳定解与由文献 [26] 给出的方法计算出的不稳定解联结起来，使得在这一电流密度的很窄小范围内，一个合适的小扰动就可实现周期活动与平衡活动的转换. 在这一范围内周期活动的消失已在 [27,28] 中用数值和实验方法给出. 因为在 I_1 出现分枝，这特殊行为是次临界的：在 I_1 出现的周期解是非稳的. 这些非稳周期解提供了一条通向大周期解的道路，而在 I_1 处的次临界分枝点就说明了不断激发的阈值的存在.

若 H–H 系统中的 $I = 0$ 和 $\overline{g}_{\text{K}}$ 从其标准值 $3.6\text{ms}\cdot\text{cm}^{-2}$ 开始减少，则在任意 $\overline{g}_{\text{K}}$ 有唯一的平衡解，且该解在次 Hopf 分枝点接近 3.84 和 $19.26\text{ms}\cdot\text{cm}^{-2}$ 时改变其稳定性. $\overline{g}_{\text{K}}$ 的变化可解释为膜通道密度的变化，其产生原因是由于通道产生和消失的比率失调或者是由于膜域的改变，并且该改变可以用四乙铵 (TEA) 或 4- 氨基吡啶 (4AP) 阻塞 K^+ 选择通道的方法来模拟.

因此，当标准的 H–H 方程的外加电流密度 I 或在 $I = 0$ 时 H–H 方程的最大 K^+ 的导电率 $\overline{g}_{\text{K}}$ 有一个发生变化时 (作为分枝参数)，平衡解的唯一分枝点趋向于 Hopf 分枝点的小振幅的周期解. 然而当 $(\overline{g}_{\text{K}}, I)$ 改变时，多平衡点发生在较小的 $\overline{g}_{\text{K}}$ 和超极化电流密度处 (见图 3.6.4).

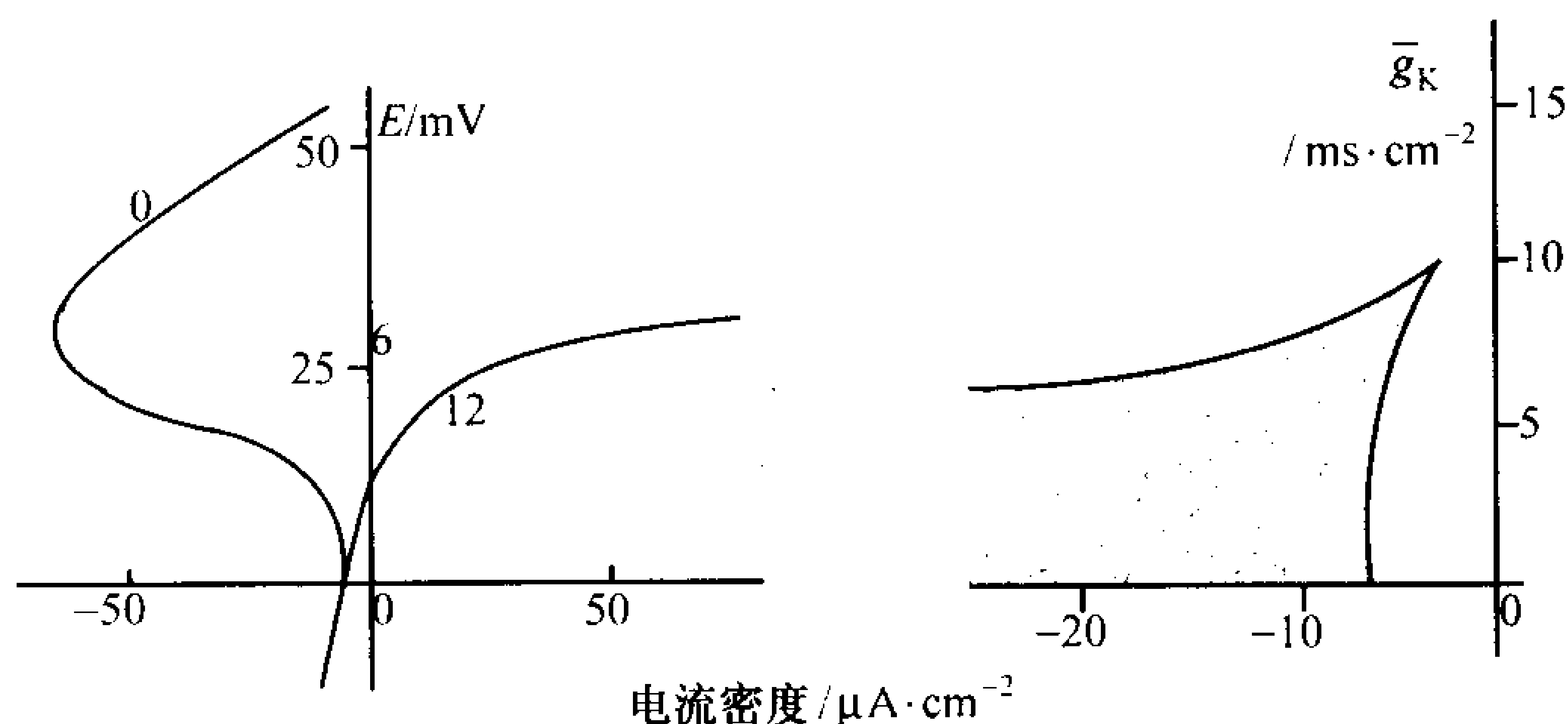


图 3.6.4　H–H 方程在小的 g_{K} 与超极化电流密度处平衡点解

当 I 或 $\bar{g}_K$ 作为分枝参数时，分枝发生在点处. 当 $(\bar{g}_K, I)$ 是分枝参数，分枝沿 $\bar{g}_K$–I 平面的曲线出现. 图 3.6.5 给出了平面 $\bar{g}_K$–I 上的 Hopf 分枝点，坐标轴是标准的 $\bar{g}_K$ ($36\mathrm{ms\cdot cm^{-2}}$) 和零电流.

$\bar{g}_K$ 取标准值，去极化电流在先经过此临界值再经过超临界值 Hopf 分枝时递增，电流为零时，$\bar{g}_K$ 在经过次 Hopf 分枝点时发生递减，因此，当有一个沿分枝曲线发生变化时，次分枝与超分枝间也发生变动，这些变化在曲线的实心点处.

实心点是曲线的终点，单个特征值的复共轭对的实部与虚部全为零，或分枝发生在平衡点处. 从终点向上，曲线连续地穿越 $\bar{g}_K$ 轴 (在一个较小的去极化电压下)，然后穿越 I 轴，在图 3.6.3 的 I_1 点，该曲线是一个次临界分枝点的轨迹. 在空心点处分枝点变成了超临界点，曲线在图 3.6.3 的 I_2 点超过或穿越 $\bar{g}_K$ 标准值，然后在空心点返回一个次临界分枝点，该点超过其另一分枝且在 $\bar{g}_K$ 为零时才终止，因此图 3.6.1 中最大和最小平衡点均产生分枝点.

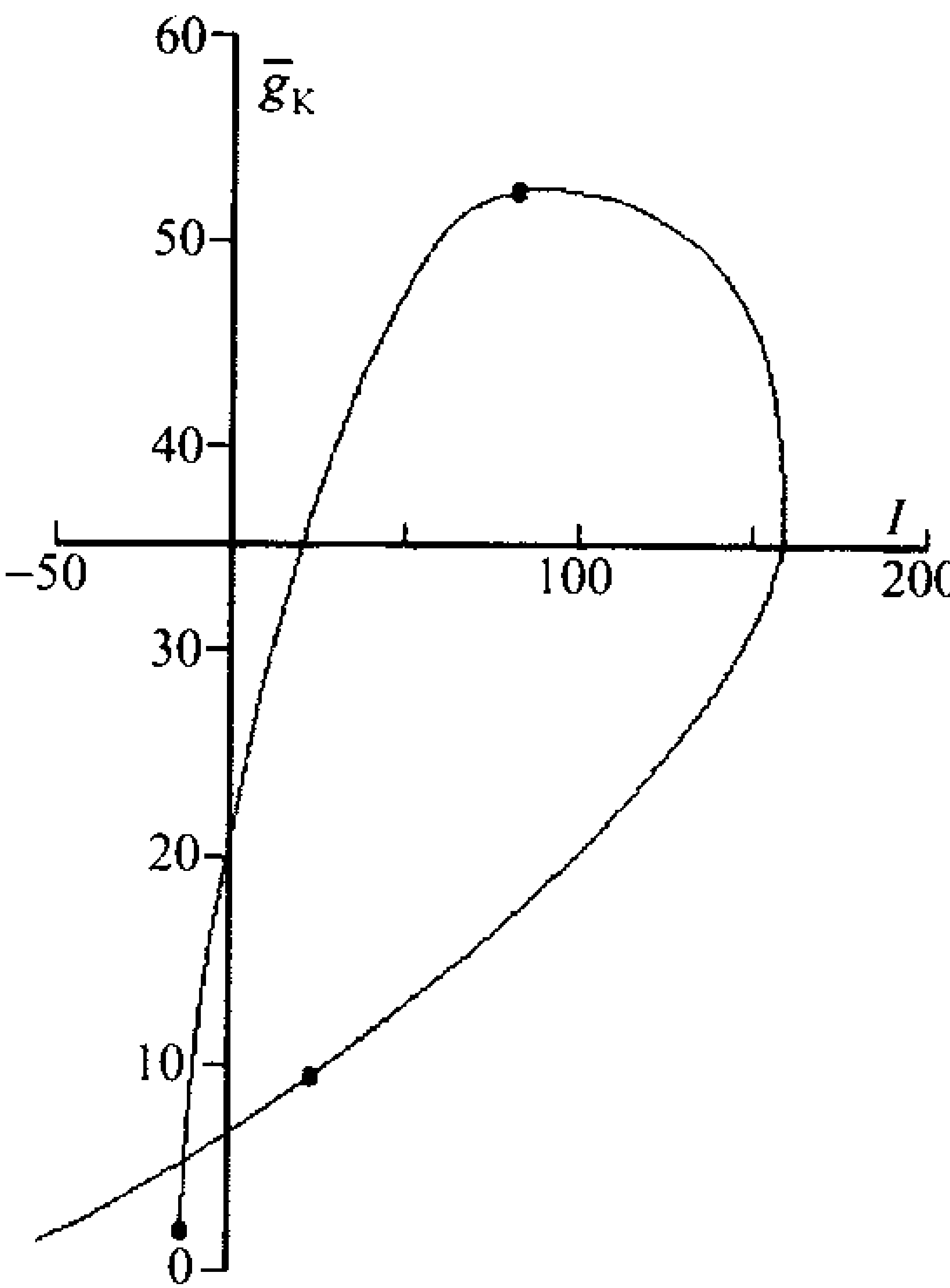


图 3.6.5　H–H 方程在 $\bar{g}_K$–I 平面中的 Hopf 分枝曲线

注意到对于 H–H 方程，人们仅仅发现 Hopf 分枝进入周期，没有发现非周期到周期分枝. H–H 方程表现为二维特征：事实上

(V,m,n,h) 可简化成带有 $m-m_\infty$ 和 $h-(1-n)$ 类似性质的 (V,n) 系统.

3.6.5 大振幅周期解

如果我们考察 3 个分枝参数, Hopf 分枝曲面将把稳定平衡解与小振幅周期解分隔出来. Hopf 分枝理论仅能描述解行为接近分枝曲面的情况. 然而, 猜想分枝曲面形成了参数空间的一个闭区域似乎是合理的, 这使得一个大幅度周期解与作用电势的不断放电相对应. 在一个超临界分枝点附近, 随着与分枝点距离的增加, 小幅度稳定周期解成为大幅度周期解 (如图 3.6.3(b)), 在次临界分枝点附近, 能跳到一个稳定大振幅周期解的一个分枝. 这里我们用几何方法来建立改变 $\bar{g}_\mathrm{K}$ 与外加电流时 H–H 方程的大振幅周期解的存在性理论作为 Holden 和 Yoda 的理论的推广.

H–H 系统包含两个快速 (V,m) 过程和两个慢过程 (n,h): 在时间度量上的差异首次由 EitzHugh 在研究 H–H 系统的阀值特征时发现. 引进新时间度量 $\tau=\varepsilon t$, 其中 ε 很小, H–H 系统可以写成

$$\begin{cases} \varepsilon\dot{V}=-I-F(V,m,n,h), \\ \varepsilon\dot{m}=\lambda_m(V)(m_\infty(V)-m), \\ \dot{n}=\lambda_n(V)(n_\infty(V)-n), \\ \dot{h}=\lambda_h(V)(h_\infty(V)-h), \end{cases} \tag{3.6.5}$$

这里 “·” 表示关于 τ 的微分, $m_\infty(V)$ 是一个递减的 S 型曲线 (一个有界单调仅有一个拐点的函数), 而 $n_\infty(V)$ 和 $h_\infty(V)$ 是一个递增的 S 型函数.

对于 $F(V,m,n,h)$ 假定其形式为

$$F=g_1(m)g_2(h)(V-V_\mathrm{Na})+g_\mathrm{K}(n)(V-V_\mathrm{K})+g_\mathrm{L}(V-V_\mathrm{L}),$$

这里 V_Na, V_K 和 V_L 是钠、钾和泄漏的平衡点电压值, 乘积 $g_1(m)\cdot$

$g_1(h)$ 是钠电导率，g_L 是一个常值漏导率.

当令 $\varepsilon=0$ 时，便得到其退化系统

$$\begin{cases} I+F(V,m,n,h)=0, \\ \lambda_m(V)(m_\infty(V)-m)=0, \\ \dot{n}=\gamma_n(V)(n_\infty(V)-n), \\ \dot{h}=\gamma_h(V)(h_\infty(V)-h). \end{cases} \tag{3.6.6}$$

退化系统的解一定在全系统的 V, m 方程所定义的二维子流形上，并且四维全系统因此而被简化为一个二维系统，该系统的解很容易描述，因为 ε 很小，所以全系统的任意解预计与退化系统相应的解很近. 我们先分析退化系统，然后分析全系统.

1. 退化系统

首先我们确定由 V, m 方程 (3.6.6) 所定义的二维子流形系统的特征. 由第一个方程解出 V 为

$$V=V_{\mathrm{Na}}+\frac{k(n,h)}{g_1(m)+C(n,h)},$$

这里

$$k(n,h)=\frac{g_{\mathrm{K}}(n)(V_{\mathrm{K}}-V_{\mathrm{Na}})+g_{\mathrm{L}}(V_{\mathrm{L}}(V_{\mathrm{L}}-V_{\mathrm{K}}))}{g_2(h)},$$

其中 $\dfrac{\partial k}{\partial n}>0$, $\dfrac{\partial k}{\partial h}<0$, 并且 $C(n,h)=\dfrac{g_K(n)+g_L}{g_2(h)}$, $\dfrac{\partial C}{\partial n}>0$, $\dfrac{\partial C}{\partial h}<0$. 假定对任意 $C>0$, $\dfrac{1}{C+g_1(m)}$ 是一个 S 形函数，其中 m 满足 $m=m^x$, $x>1$. 图 3.6.6 给出了 $V(m,n,h)$ 和 $m_\infty(V)$ 的图形的 3 种相交情况.

(a) 当 h 很小时，

$$V=V_{\mathrm{K}}+\frac{g_{\mathrm{L}}(V_{\mathrm{L}}-V_{\mathrm{K}})}{g_{\mathrm{K}}(n)+g_{\mathrm{L}}},$$

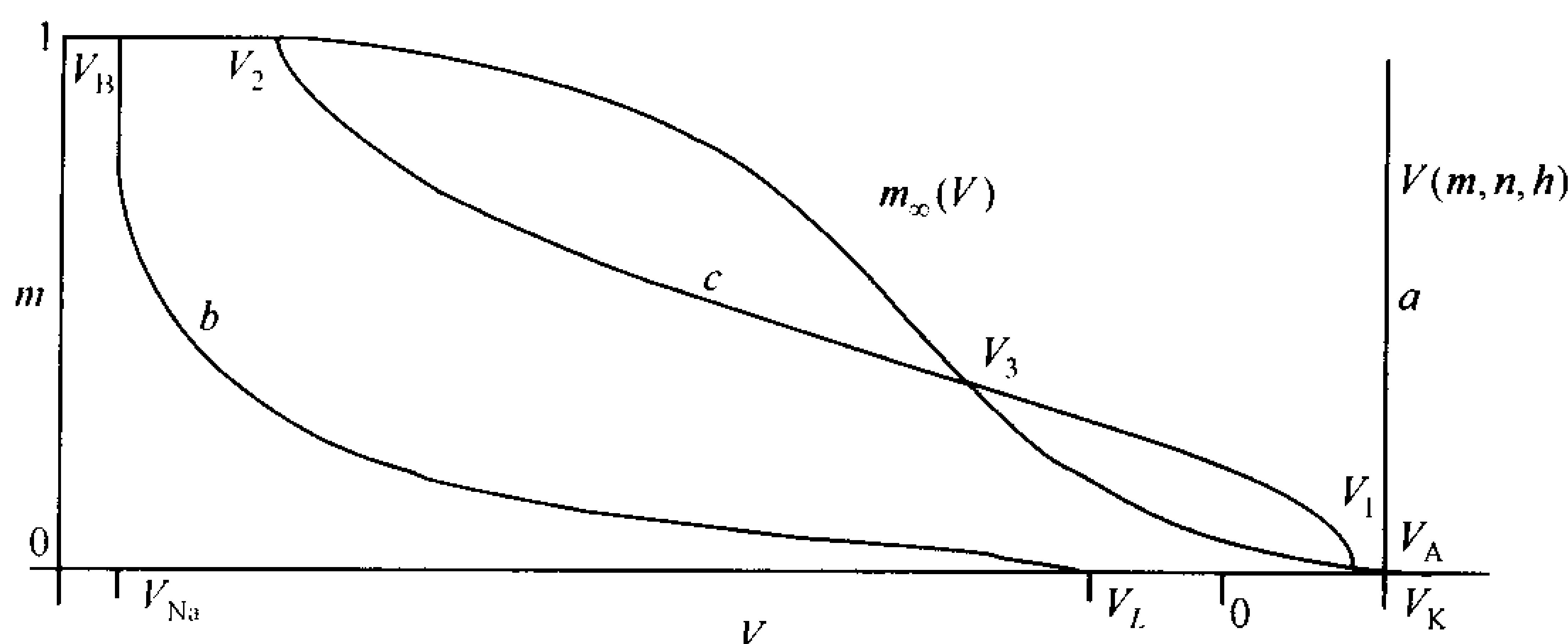


图 3.6.6　$V(m,n,h)$ 同 $m_\infty(V)$ 图形相交的 3 种情况 a, b, c

V 几乎独立于 m, 当 $h=0$ 时便完全独立于 m, 且一旦 V_A 接近 $V_{\rm K}$, $V(m,n,h)$ 与 $m_\infty(V)$ 便相交.

(b) 当 C 很小时,

$$V = kg_1(m),$$

一旦 V_B 与 $V_{\rm Na}$ 接近时, $V(m,n,h)$ 与 $m_\infty(V)$ 便相交.

(c) 对于 h 和 k 都不太小的中间情形. $V(m,n,h)$ 和 $m_\infty(V)$ 相交 3 次于 v_1, v_2, v_3. 图 3.6.7 给出了 V-m 平面内的向量场, 很明显, 当 (m,h) 固定时, v_1, v_2 是稳定的, 而 v_3 是非稳定的.

情形 (a) 当 $h=0$ 时必发生, 而情形 (b) 或 (c) 是否发生决定于参数的选取. 假定所有情形都发生, 且固定 n 为 n_0, h 从 0 变化到 1. 当 $h=h_1(n_0)$ 时, 出现从情形 (a) 到 (c) 的变化, 当 $h=h_2(n_0)$ 时出现从情形 (c) 到 (b) 的变化 (见图 3.6.8).

在区域 $D_{\rm I}=\{(n,h)\mid 0<n,h<1,h_1<h_2(n)\}$ 中可定义光滑函数 $V_{\rm I}=v_1$, $M_{\rm I}=m_\infty(V_{\rm I})$. 同样地, 在 $D_{\rm II}=\{(n,h)\mid 0<n,h<1,h>h_2(n_0)\}$ 上定义光滑函数 $V_{\rm II}=v_2$, $M_{\rm II}=m_\infty(v_2)$. 而在 $D_{\rm III}=D_{\rm I}\cap D_{\rm II}$ 上定义 $V_{\rm III}=v_3$, $M_{\rm III}=m_\infty(v_3)$. h_i $(i=1,2)$ 由下面的方程定义

$$V(m,n,h)-U(m)=0,\quad \frac{\partial V}{\partial m}-\frac{{\rm d}U}{{\rm d}m}=0, \tag{3.6.7}$$

图 3.6.7 $V(m,n,h)$ 与 $m_\infty(V)$ 相交的 3 种情况的向量域

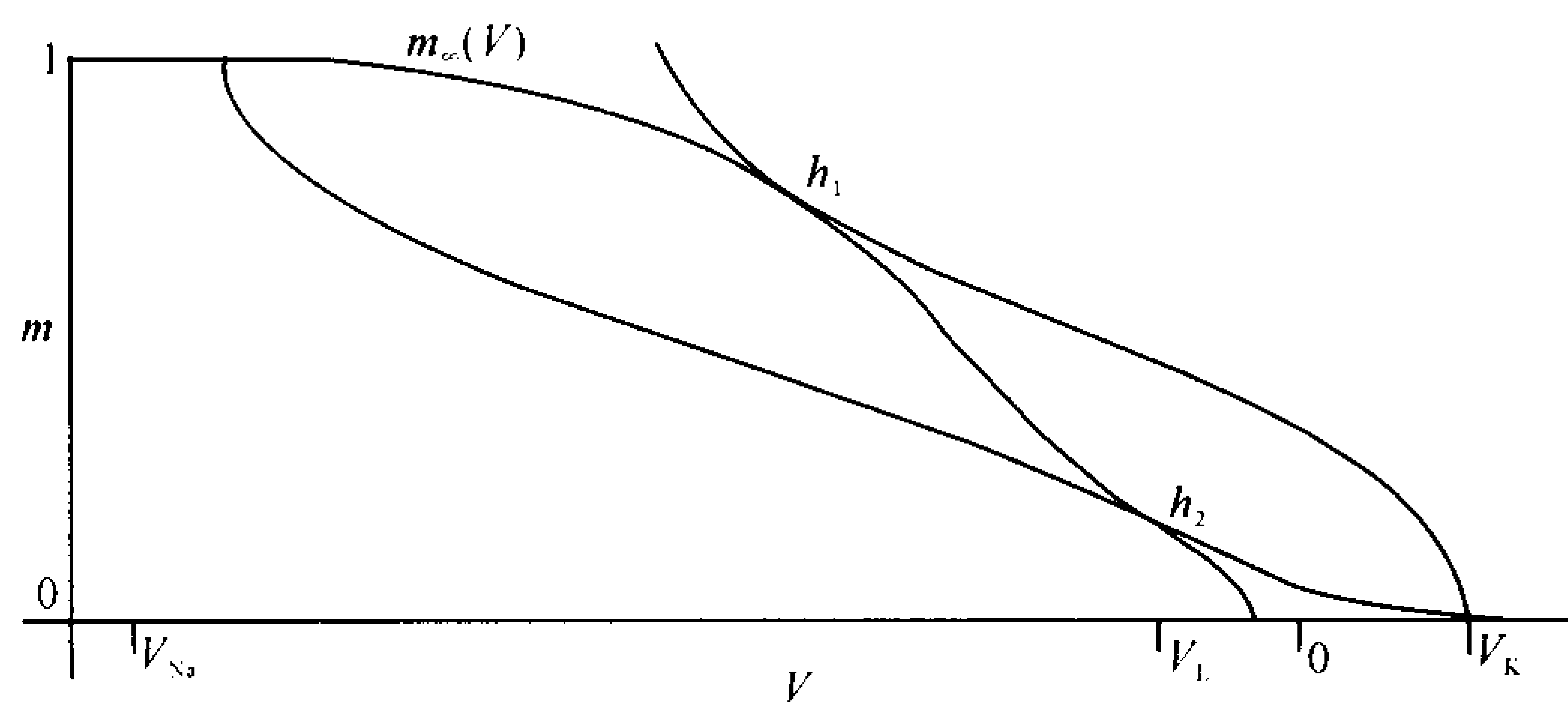


图 3.6.8

这里 $U(m)$ 是 $m_\infty(V)$ 的一个单调递增函数.

方程 (3.6.7) 的 Jacobi 矩阵是

$$\begin{pmatrix} \dfrac{\partial V}{\partial m}-\dfrac{\mathrm{d}U}{\mathrm{d}m} & \dfrac{\partial V}{\partial h} \\ \\ \dfrac{\partial^2 V}{\partial m^2}-\dfrac{\mathrm{d}^2U}{\mathrm{d}m^2} & \dfrac{\partial^2 V}{\partial m\partial h} \end{pmatrix}.$$

很明显 $\dfrac{\partial V}{\partial h}<0$ 且 $\dfrac{\partial^2 h}{\partial m^2}-\dfrac{\mathrm{d}^2u}{\mathrm{d}m^2}$ 在 h_1 处不小于 0, 而在 h_2 处不大于 0.

假定在 h_i $(i=1,2)$ 处严格不等式成立, h_i 可表示为 n 的函数, 图 3.6.9 给出了 H–H 方程的 h_i 的数值计算结果.

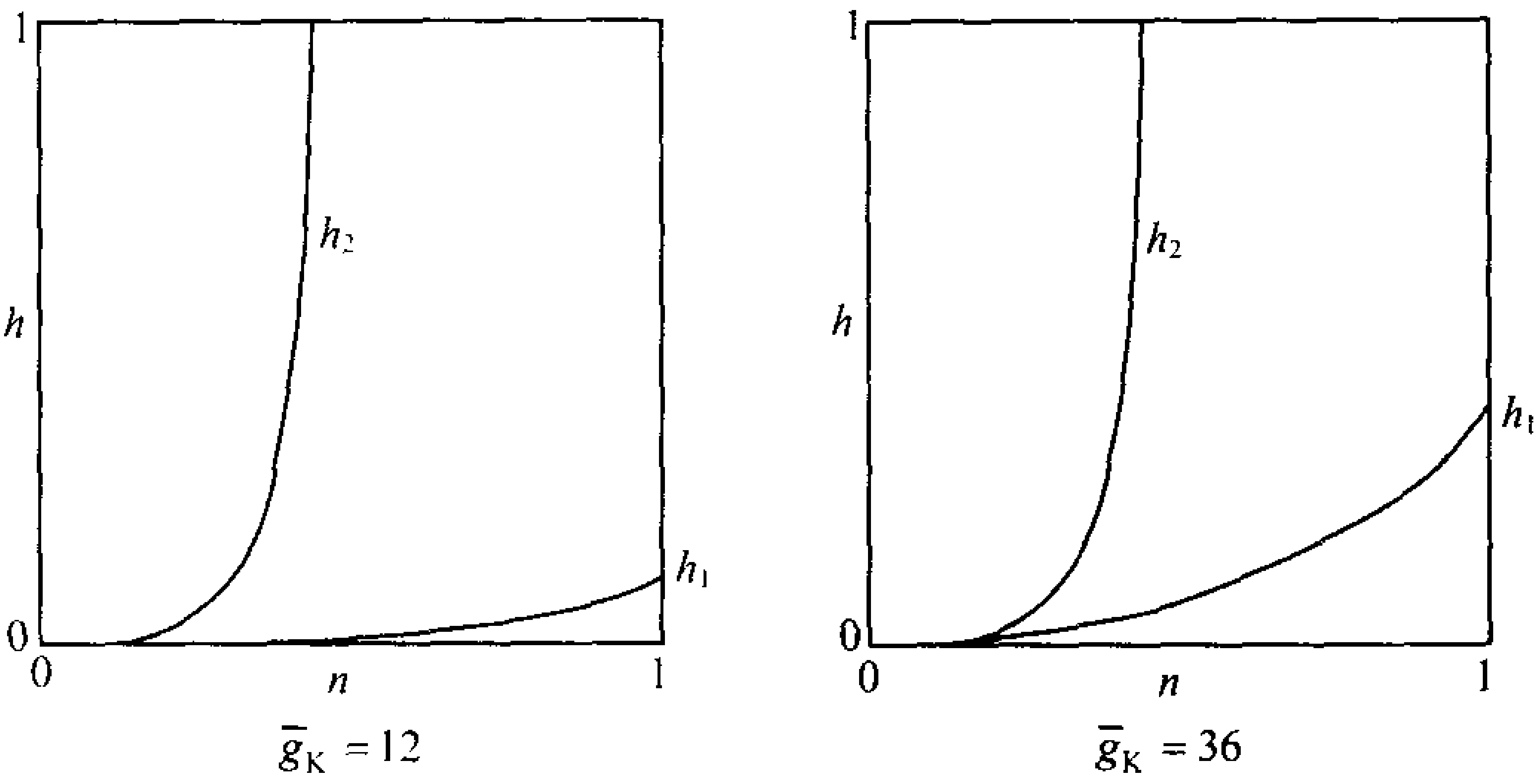


图 3.6.9　H–H 方程 h_1, h_2 的数值计算结果

这样我们知道在 D_{I} 内二维流形可以表示为 $S_{\mathrm{I}}=(V_{\mathrm{I}}(n,h), m_\infty(V_1))$, 在 D_{II} 内为 $S_{\mathrm{II}}=(V_{\mathrm{II}}(n,h),m_\infty(V_{\mathrm{II}}))$, 在 $D_{\mathrm{III}}=D_{\mathrm{I}}\cap D_{\mathrm{II}}$ 内有新的分枝 $S_{\mathrm{III}}=(V_{\mathrm{III}}(n,h), m_\infty(V_{\mathrm{III}}))$ 出现.

既然在 n, h 均固定的系统中 $(V_{\mathrm{I}},m_\infty(V_{\mathrm{I}}))$ 和 $(V_{\mathrm{II}},m_\infty(V_{\mathrm{II}}))$ 是稳定的, 而 $(V_{\mathrm{III}}, m_\infty(V_{\mathrm{III}}))$ 是非稳定的, 则退化系统 (3.6.6) 的任意解或在 S_{I} 上或在 S_{II} 上.

设 Γ 是关于 $(V,m_\infty(V),n_\infty(V),h_\infty(V))$ 上的一幅图, 则有下面 4 种可能:

(i) $S_{\rm I}$ 与 $S_{\rm II}$ 均与 Γ 相交;

(ii) $S_{\rm I}$ 与 Γ 相交但 $S_{\rm II}$ 与 Γ 不相交;

(iii) $S_{\rm II}$ 与 Γ 相交但 $S_{\rm I}$ 与 Γ 不相交;

(iv) $S_{\rm I}$ 和 $S_{\rm II}$ 都不与 Γ 相交.

我们对情形 (i) 和 (ii) 不感兴趣. 在情形 (iii) 中, 假定初始条件 Q 不在 $S_{\rm I}$ 和 $S_{\rm II}$ 上: 轨线是一条 n–h 平面上的瞬时直线, 其终点在 $S_{\rm I}$ 或 $S_{\rm II}$. 假定 Q 在 $S_{\rm II}$ 上: 则轨线是在 n–h 平面内的一条曲线的提升. 它满足

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \lambda_n(V_{\rm II}(n,h))(n_\infty V_{\rm II}(n,h) - n),$$
$$\frac{\mathrm{d}h}{\mathrm{d}t} = \lambda_h(V_{\rm II}(n,h))(n_\infty V_{\rm II}(n,h) - h).$$

因为 Γ 与 $S_{\rm II}$ 不相交, 在有限时间内, 轨线终止于 h_1. 假定 Q 在 h_1 上: 轨线是一条 n–h 平面上的瞬时直线, 其终点在 $S_{\rm I}$ 上. 设 Q 在 $S_{\rm I}$ 上, 轨线是 n–h 平面上的一条曲线的提升, 并满足

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \lambda_n(V_{\rm I}(n,h))(n_\infty V_{\rm I}(n,h) - n),$$
$$\frac{\mathrm{d}h}{\mathrm{d}t} = \lambda_h(V_{\rm I}(n,h))(h_\infty V_{\rm I}(n,h) - h).$$

因为 Γ 与 $V_{\rm I}$ 相交, 所以曲线收敛于交点. 假定 Q 在 h_2 上, 则轨线为一条 n–h 平面上的瞬时直线并在 $S_{\rm II}$ 上终止. 图 3.6.10(a) 给出了整条轨线.

对情形 (iv), 其轨线除 Q 在 $S_{\rm I}$ 上以外与情形 (iii) 相同. 因为 $S_{\rm I}$ 与 Γ 不相交, 曲线不收敛到任何点且在有限时间内达到 h_2. 图 3.6.10(b) 给出了整条轨线. 在此情形的任意轨线从 h_2 上的 C 开始返回到某些 h_2 上的 C^*. 定义 h_2 到自身的函数为 $\phi(C) = C^*$, ϕ 是一个连续函数, 则由 Brouwer 不动点原理保证存在一个不动点 $\overline{C}$, 使得 $\phi(\overline{C}) = \overline{C}$, 这意味着退化系统的周期解 Λ 是存在的.

为了讨论周期解 Λ 的稳定性, 我们假定 $V_{\rm I}$ 和 $V_{\rm II}$ 与 $V_{\rm K}$ 和 $V_{\rm Na}$ 分布很接近, 这样 $\left.\dfrac{\mathrm{d}n_\infty}{\mathrm{d}V}\right|_{V\in V_{\rm II}}$ 和 $\left.\dfrac{\mathrm{d}h_\infty}{\mathrm{d}V}\right|_{V\in V_{\rm II}}$ 很小且 $\dfrac{\mathrm{d}\lambda_n}{\mathrm{d}n}$, $\dfrac{\mathrm{d}\lambda_h}{\mathrm{d}V}$ 也

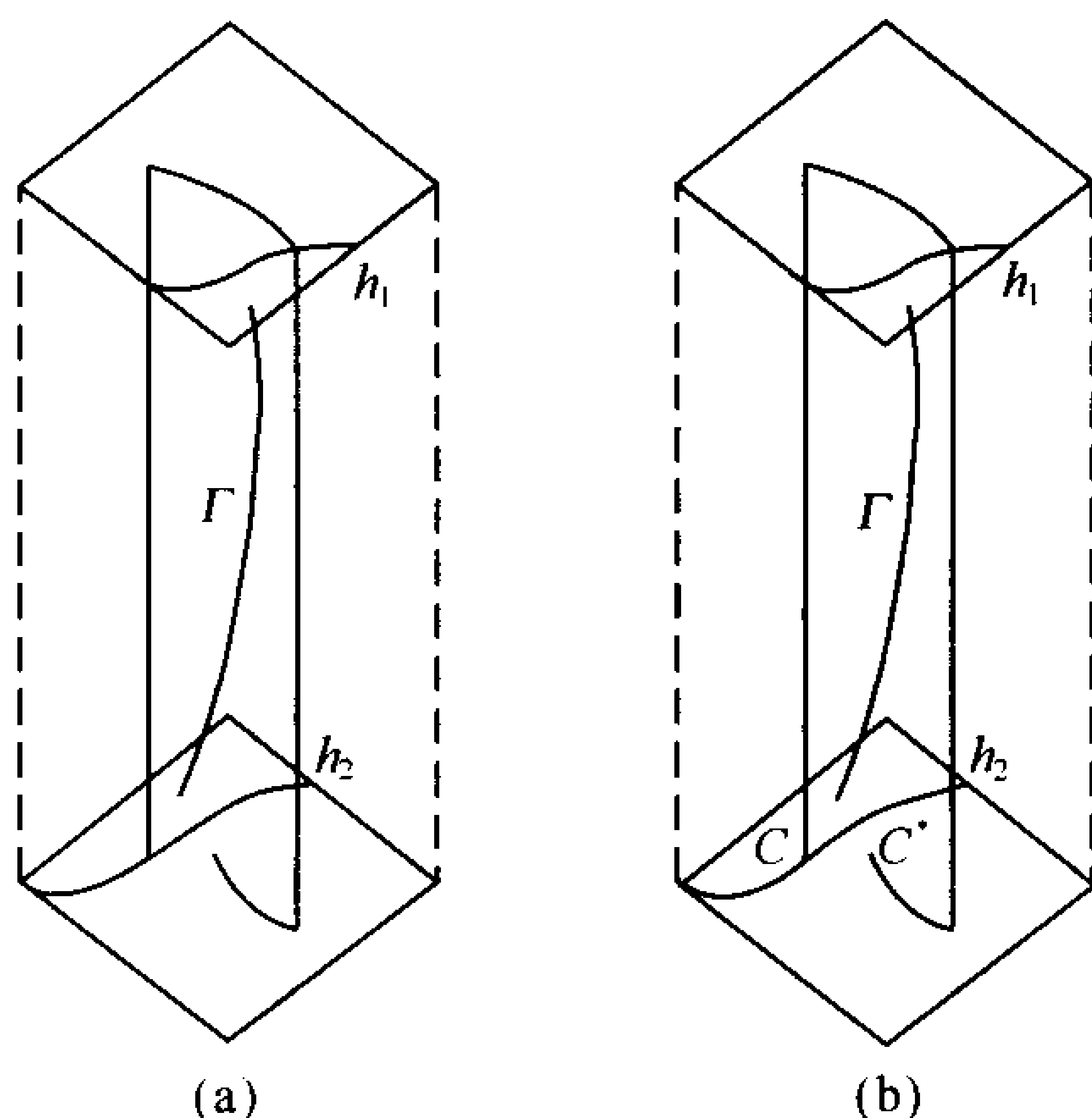


图 3.6.10

很小. 周期解 Λ 的稳定性可以由讨论变分方程来确定. 设 δn, δh 是与周期解 Λ 的一个小偏差, δn, δh 满足线性方程组

$$\begin{aligned}\delta\dot{n} &= \gamma_h + \frac{\mathrm{d}\gamma_n}{\mathrm{d}V}\frac{\partial V}{\partial n}(n_\infty - n) + \gamma_n\frac{\mathrm{d}\gamma_\infty}{\mathrm{d}V}\frac{\partial V}{\partial n}\delta n\\ &\quad+\frac{\mathrm{d}\gamma_n}{\mathrm{d}V}\frac{\partial V}{\partial n}(n_\infty - n) + \gamma_n\frac{\mathrm{d}\gamma_n}{\mathrm{d}V}\frac{\partial V}{\partial n}\delta h,\\ \delta\dot{h} &= \frac{\mathrm{d}\gamma_n}{\mathrm{d}V}\frac{\partial V}{\partial n}(h_\infty - h) + \gamma_n\frac{\mathrm{d}\gamma_\infty}{\mathrm{d}V}\frac{\partial V}{\partial n}\delta n\\ &\quad+\gamma_h + \frac{\mathrm{d}\gamma_h}{\mathrm{d}V}\frac{\partial V}{\partial h}(h_\infty - h) + \gamma_h\frac{\mathrm{d}\gamma_\infty}{\mathrm{d}V}\frac{\partial V}{\partial n}\delta h,\end{aligned}\tag{3.6.8}$$

可以在 Λ 的慢相处估计方程中每个导数, 因为假定 $\dfrac{\mathrm{d}\gamma_n}{\mathrm{d}V}$, $\dfrac{\partial m_\infty}{\partial V}$, $\dfrac{\mathrm{d}\gamma_h}{\mathrm{d}V}$, $\dfrac{\partial h_\infty}{\partial V}$ 在慢相位时很小, 且其他所有项都是有界的, 从而得到

$$\begin{aligned}\delta\dot{n} &= (-\gamma_n + \varepsilon_1)\delta n + \varepsilon_2\delta h,\\ \delta\dot{h} &= \varepsilon_3\delta n + (-\gamma_h + \varepsilon_4)\delta h.\end{aligned}$$

对任意 $l > 0$, 正方形 $\{(\delta n, \delta h) \mid |\delta n| \leqslant l, |\delta h| \leqslant l\}$ 是系统的正的不变集, 这是保证周期解 Λ 稳定的充分条件.

2. 完全系统周期解的存在性

将具有初始条件 Q 的全系统周期解记为 $X_\varepsilon(Q)$. 我们需要如下引理:

引理 3.6.1 在相空间的任何有界部分, 曲线 $X_0(Q)$ 可作为 $X_\varepsilon(Q)$ 的第零次渐近近似, 即当 $\varepsilon \to 0$ 时, $X_\varepsilon(Q) \to X_0(Q)$, 在有限长度的 $X_0(Q)$ 的任意一段区间上一致收敛.

该引理的证明由 Mishchenko 和 Rozor 给出.

取一个小的线段 P, 该区间横卧在周期解 Λ 的慢分枝, 并取一个二维的圆盘 D, 该圆盘在 V–m 平面上其半径不依赖于 ε. 因为 Λ 是稳定的, 由引理 3.6.1, 当 ε 很小时, 从 $P \times D$ 出发的任意轨线必在有限时间内返回, 因此 Brouwer 不动点定理给出了完全系统周期解的存在性 (见图 3.6.11).

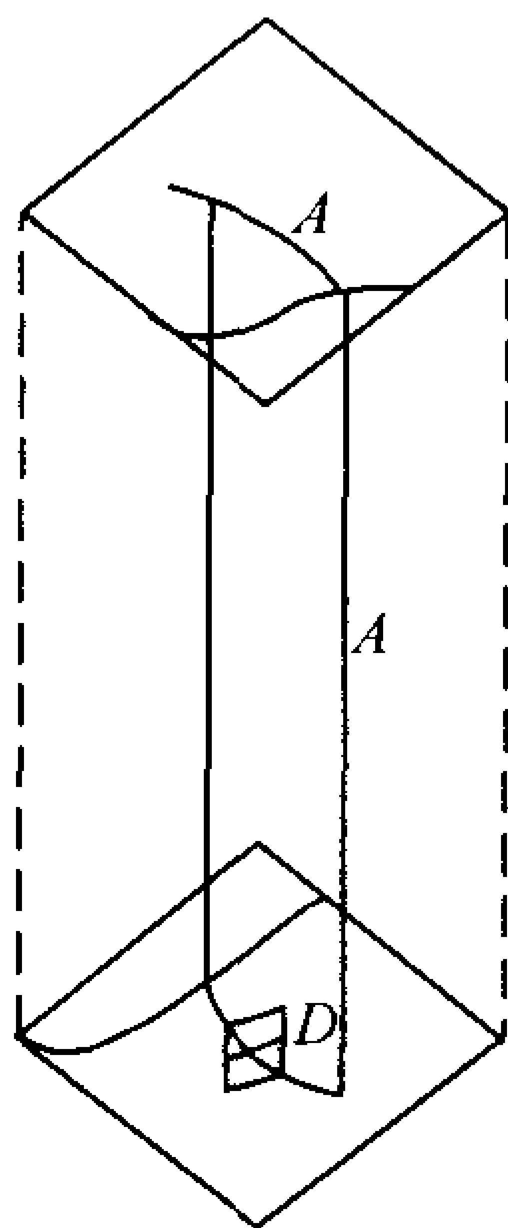


图 3.6.11

3.6.6 周期解间的分枝

1. 倍周期

有强迫项的非线性微分方程是重要的方程, 其解关于时间 t 是周期的. 要理解 R^n 空间 ($n \geqslant 2$) 中的**倍周期**, 具有基本的 Floquet

理论是有益的. **Floquet 指数** 是通过关系 $\lambda(T)=\exp(\sigma T)$ 定义为 $\sigma=\xi+\mathrm{i}\eta$ 的, 这里 T 是周期. $\lambda(T)$ 是 **Floquet 乘子**, 它表示单位圆. 当乘子的共轭复数对 $\lambda, \overline{\lambda}$ 离开单位圆 (见图 3.6.12) 时, 周期解失稳.

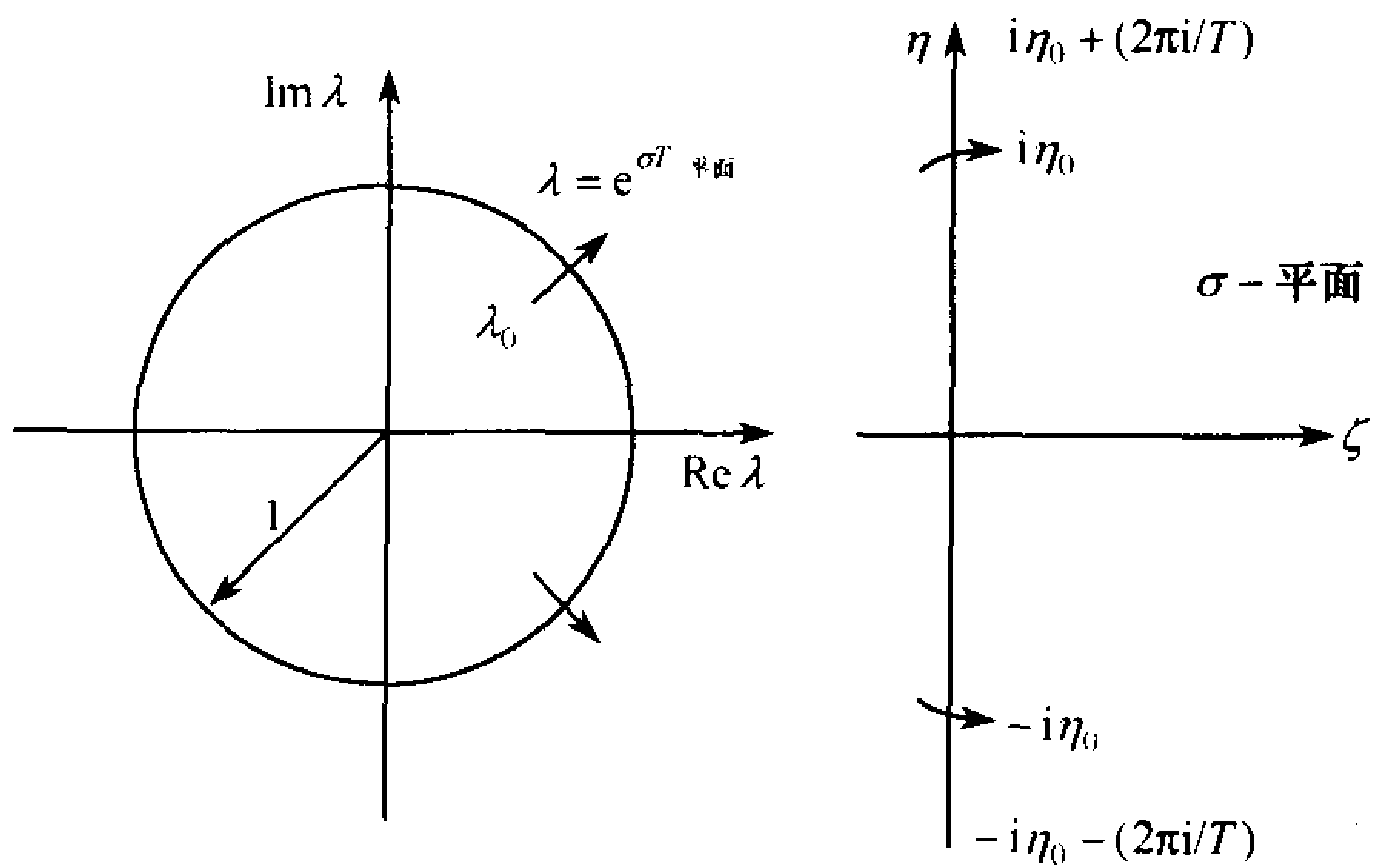


图 3.6.12

神经元经常隐含着周期模式, 而不是简单周期的放电: 这种周期模式的活动可能是它们的正常行为方式或者可能是由不正常的环境导致的. 对每一种情形, 其行为是参数值变动的结果, 因此病态行为可以认为是通过一系列的分枝后, 正常行为随参数变化的结果.

神经元内的复杂活动可以分为简单模式和阵发模式. 图 3.6.13 给出了一个简单模式的例子: 其活动是一系列作用电压, 该电压出

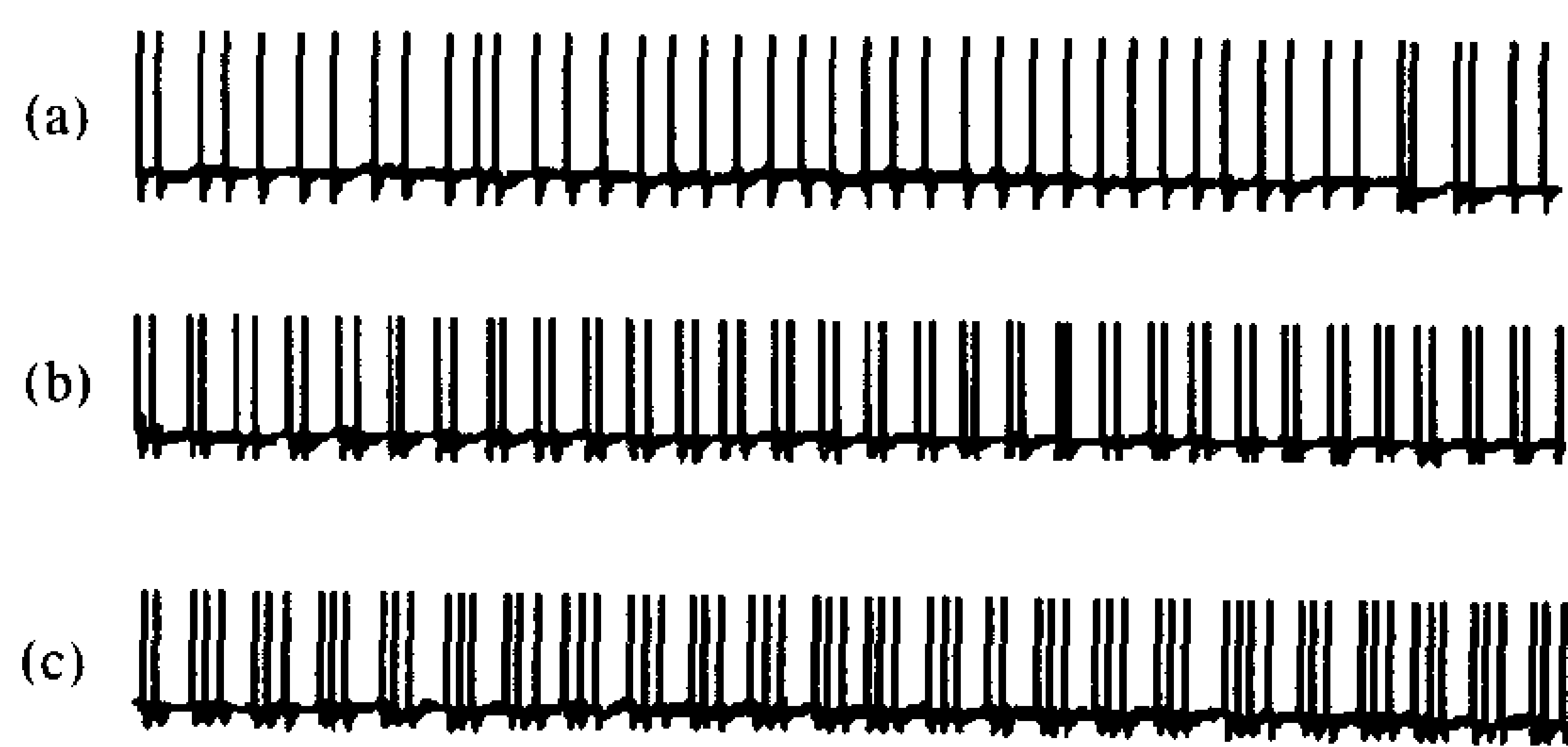


图 3.6.13

现的次数被模仿，产生了成对出现的周期现象. 尽管 H–H 方程的数值解没有给出当某一参数 (V_K, I, $\bar{g}_K$ 或 $[Ca^{++}]_0$) 发生变化时其活动的多样性, 但是 Carpenter 给出在关于传播作用电压的一个一般化 H–H 偏微分系统中存在着这种模式和阵发活动. 我们不考察描述鱿鱼巨轴突的改进的 H–H 系统，因为它不可能产生简单的周期活动, 更不用说周期模式的活动, 要获得模式的周期性宁愿从基于描述神经躯体的兴奋方程开始研究，因为该方程可以表现模式的周期性，即 (3.6.3) 所描述的 Chay 系统. 图 3.6.14 和图 3.6.15 给出了具有简单模式周期性的数值解.

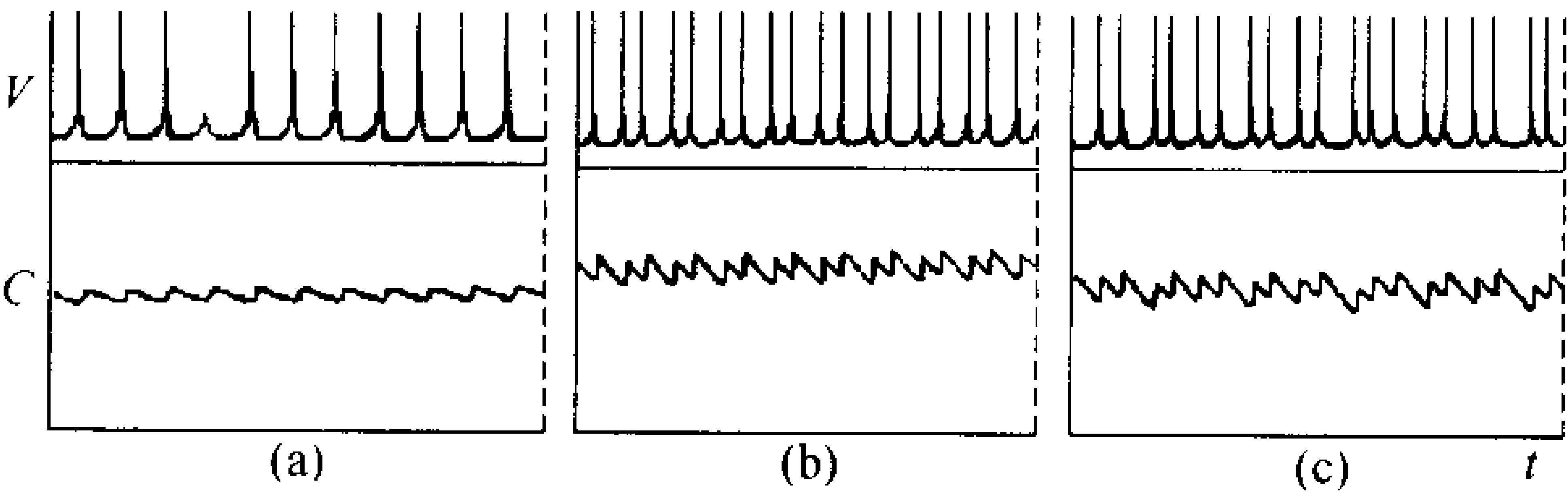


图 3.6.14

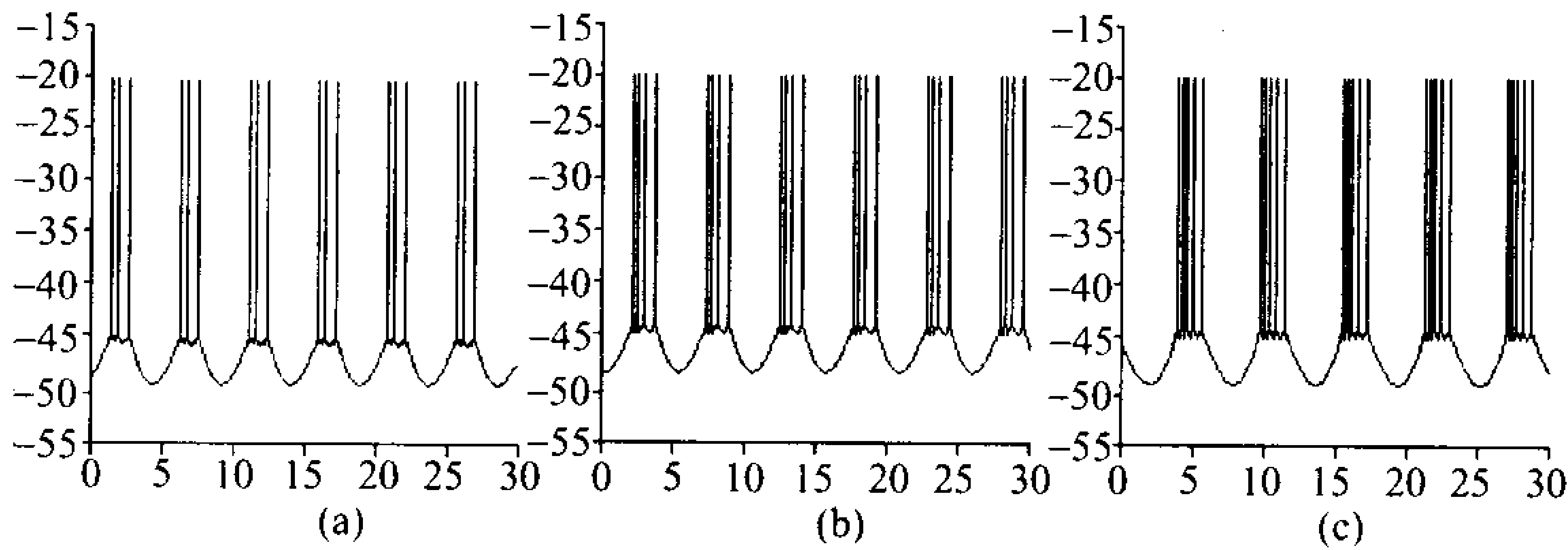


图 3.6.15

方程 (3.6.3) 中的 k_{Ca} 减少时可以观察到倍周期, 并且在 $g_{K,Ca}$–I 平面内的倍周期分叉曲线可以用 [32] 中给出的算法计算出 Floquet 指数的数值. 尽管要从 $V(t)$ 出发得到初期的倍周期为简单的模式周期是困难的，但是 $[Ca^{++}]_i(t)$ 的图形很清楚地给出了一对和两对的放电形式. 当作用电压增加时 $[Ca^{++}]_i$ 在不动点的一系列值可用来构造一个 1–D 映射 $C_{n+1} = fC_n$, 该映射是一个对称的单峰映射，当 k_{Ca} 递减时产生倍周期级联分枝： $f(C)$ 与 $C_{n+1} = C_n$

的交点是不动点，当 $\left|\frac{\mathrm{d}f}{\mathrm{d}C}\right|$ 超过 1 时变成了不稳定的不动点，并出现稳定的二周期点. 映射的迭代和 ODES 的积分给出的倍周期分枝序列相同.

2. 阵发

Carpenter 将神经元的阵发分成了 3 类：(i) 作用电压的瞬时率突然增加直至结束；(ii) 作用电压瞬时率增加在阵发中减少，就像对 Aplysia 的抛物阵发器一样；(iii) 作用电压的振幅递减成一个去极化的振动或稳定的状态，就像做了阵发性的去极化的频移. 情形 (i) 阵发可以用一个一般的 H–H 系统来产生 (但是当 H–H 方程的参数变化时的数值解不能表现出该阵发，有必要改变相关的动力学特征). 然而，情形 (ii) 阵发需要对一般化的 H–H 系统附加超导条件，由此产生了一个慢子系统，该系统对快 H–H 系统起调节作用.

图 3.6.16 给出了 Chay 系统 (3.6.3) 的阵发情况：考察当 C 的值接近在 C_H 处的一个次临界的 Hopf 分枝时的 V–n 子系统，对于 $C_H < C < Ca^{**}$, 将存在大振幅周期解，每个解都导致 C 的增加，直到 C^{**} 时阵发终止，即 V–n 子系统 (灭亡). 其行为像一个两变量振子，该振子受慢变量 C 的调节.

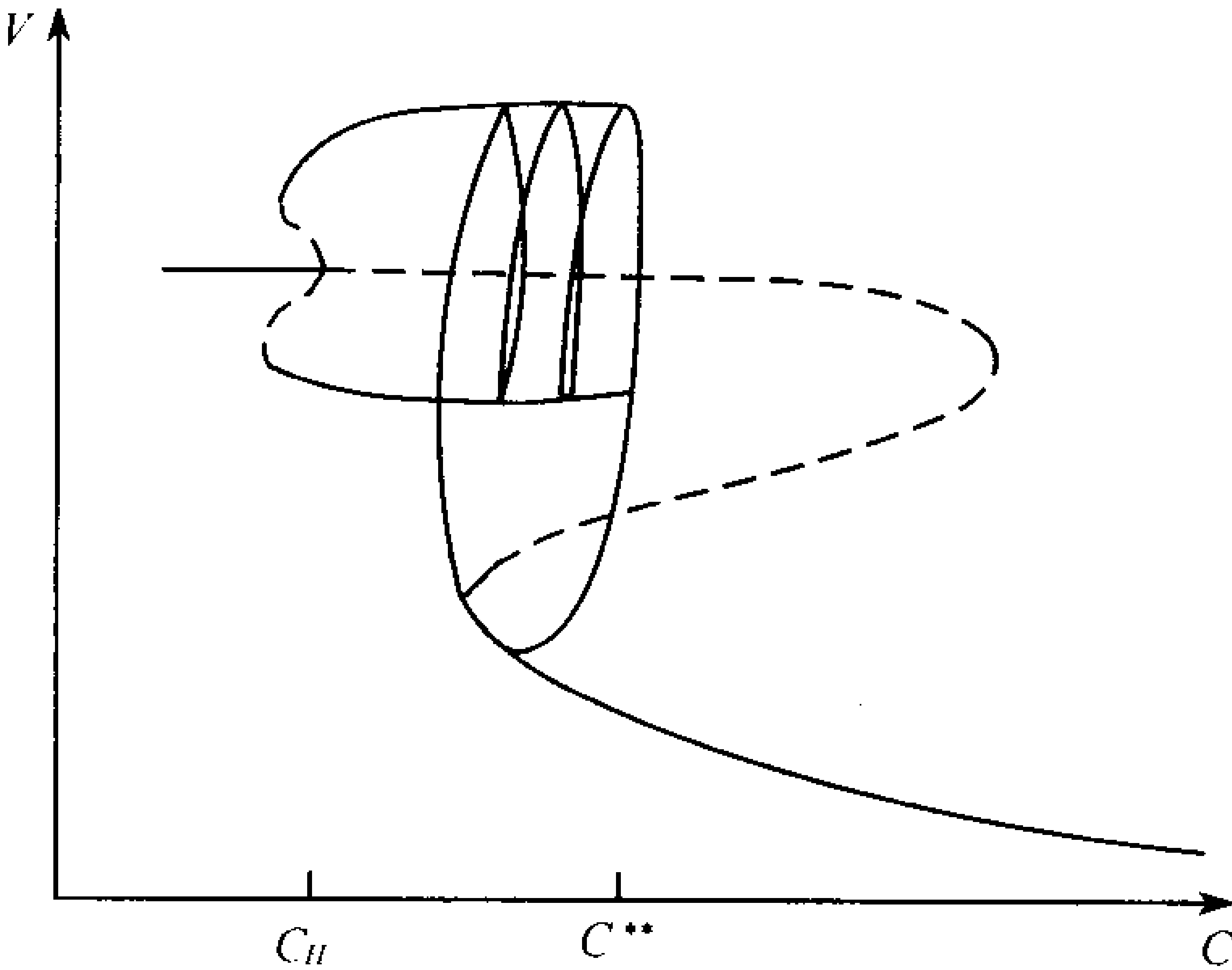


图 3.6.16

3.6.7 混沌

混沌是非线性确定性动力系统的一个复杂的非周期的暂时行为，该系统对初值高度敏感且尽管它短期是可以预测的但长期是不可预测的. 由定义知，一个动力系统就是一系列方程，其时间演化依赖于以前的行为，并可由以前的行为得到. 对于最简单的情形，系统的长期行为被限定在相空间的确定区域中，当初始条件变化不大时该区域变化也不大. 在 ω- 极限集中，这些解或者收敛于一个极限点或者在相空间中作为一个极限环进行周期振动，即轨线分别被吸引到一个不动点或周期吸引子.

然而，非线性系统也有混沌解，该解是由一个自身相当简单的确定系统产生的复杂行为的回归模型. 轨线是周期的，但吸引到一个结构相当复杂的吸引子上，该吸引子经仔细检查常表现为部分结构具有自相似性. 奇怪吸引子就位于相平面内一个复杂的超曲面上，它吸引着附近的轨线，但在吸引子中的邻近的轨线将发生指数偏离. 奇异吸引子的双曲性意味着解轨线将很快偏离，且初始条件的一个较小的不确定性随解的演化将迅速增加.

正是混沌解的这些性质使得它们难以确定且很难将其与试验数据的噪音区分开来. 然而，通过比较试验记录与实际系统的差别，一系列的几何技术已用来识别和确定混沌.

1. 用试验数据对混沌的识别

由单神经元得来的记录对混沌进行识别的技术在 [36] 中已给出讨论，在大多数的试验中仅 1 种或几种变量被记录为时间的函数. 人们通常用记录时间序列的能量谱来识别周期运动间的转换和从周期到无规律运动的转换. 能量谱不能用来区分混沌与随机运动，然而，从振动到非周期运动的路径与已知转换路径相比较可以用来确定混沌. 作为一个实例，能展现连续倍周期分枝或不规则运动的系统是一个好的选择对象，因为它能使人们相信其不规则运动就是混沌，当然另外还有许多通向混沌的道路.

能量谱的仔细研究对识别准周期是有用的，如图 3.6.17，当迫动频率与自然频率是非正常值时，准周期在周期迫动的神经元内

产生，随着作用电压振幅的慢慢调整，在正弦迫动的神经元中就出现准周期. 重构的吸引子是一个双曲环面. 能量谱在迫动频率 f_s 和自然频率 f_N 作简单叠加 $mf_s \pm nf_N$ 时达到最大.

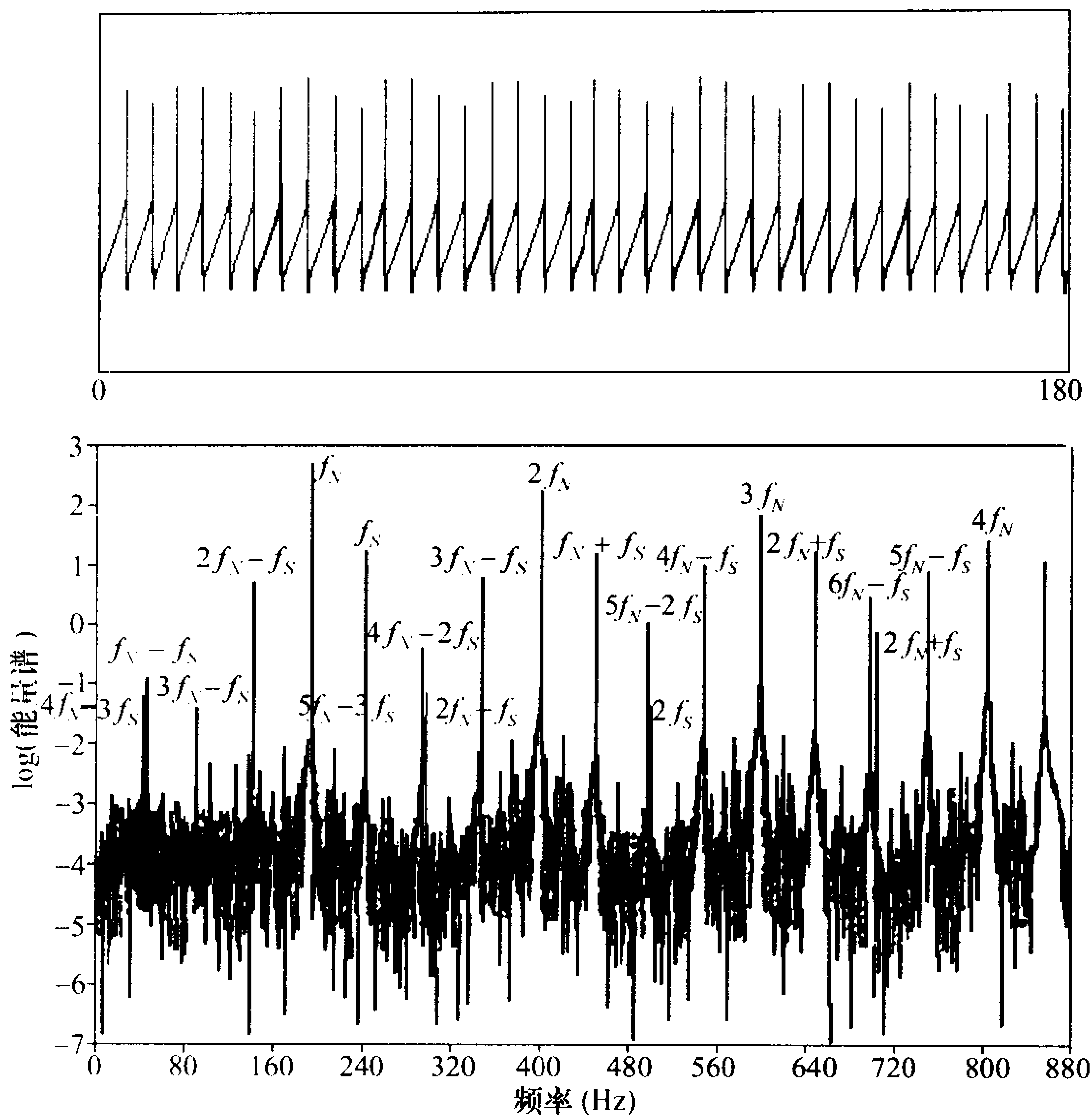


图 3.6.17

各种几何技术已发展成为能够从一个时间序列记录中抽象成为一个定量的模型. 系统的相平面图可以通过画出测定变量的微分与其自身的图形得到. Ruelle 提出用"滞后方法"从试验数据构造潜吸引子的方法. 由 Broomhead 和 King 提出的奇异值分解分析法是对吸引子构建技术的一种有价值的改进.

可以用多种方法来估计重构吸引子的特点. 我们可以用 Lyapunov 指数谱来定量描述它对初值的敏感程度，也可以估计吸引子的维数. Lyapunov 指数是无穷小体积元素的平均对数增长率，该

体积元素由相空间中的流携带，因此它使吸引子轨线的平均偏离率得到量化，利用这种方法，可以识别数值物理与生物系统中的混沌.

人们对确定神经系统中的混沌有极大的兴趣，不仅对单神经元感兴趣，而且对大神经组织也感兴趣，而更感兴趣的是关于在神经系统中混沌行为的可能意义和重要性. 尽管无规则运动是一个共同现象，几乎所有的神经活动都是不规则的，但是在研究神经元和标准神经元时发现，只有在特定条件下且对特殊小的参数范围内才可能出现混沌，这也许是因为模型比较简化的缘故，整个大脑是一个大系统网络作用.

2. 有周期强迫的兴奋系统

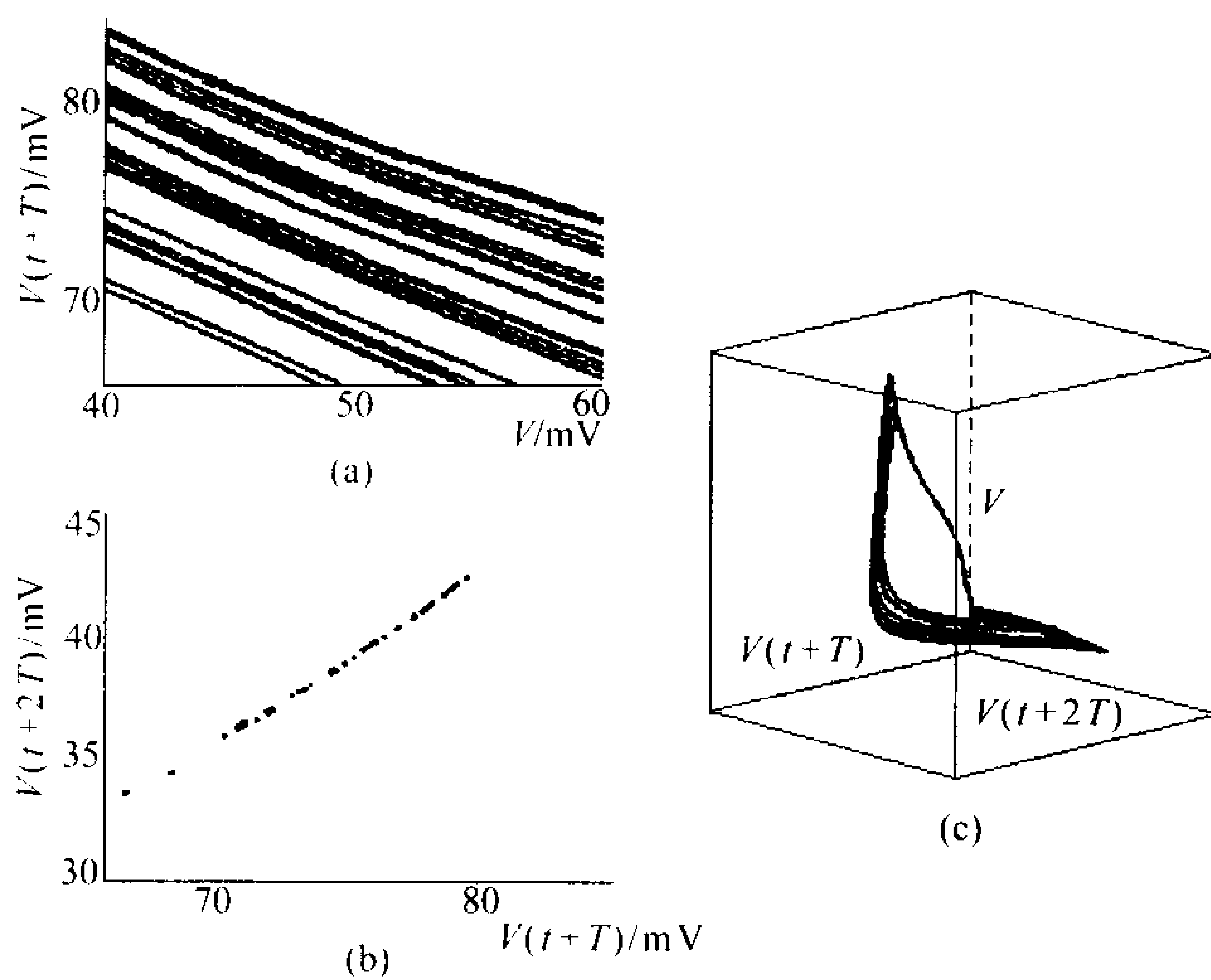
尽管 H–H 方程仅给出了平衡解或周期解，但是对于一个以合适的比率和振幅进行周期追动的非线性振子的两变量系统来说，是很容易获得混沌行为的. 对鱿鱼巨轴突膜系统，利用数值方法和试验方法，人们已研究了 H–H 膜方程对周期输入的响应情况.

相平面和状态空间的投影如图 3.6.18.

神经元对正弦电流密度 $I = A\cos(2\pi ft) + B$ 的响应包括作用电压，该电压对于驱动环是锁相的， N 作用电压在 M 环上的固定相位处出现. 对于某些频率和振幅的外加正弦电流密度， H–H 方程的解是不规律的 (见图 3.6.19). 在正弦驱动的软体动物的神经元中以及鱿鱼巨轴突中也有同样不规则的运动.

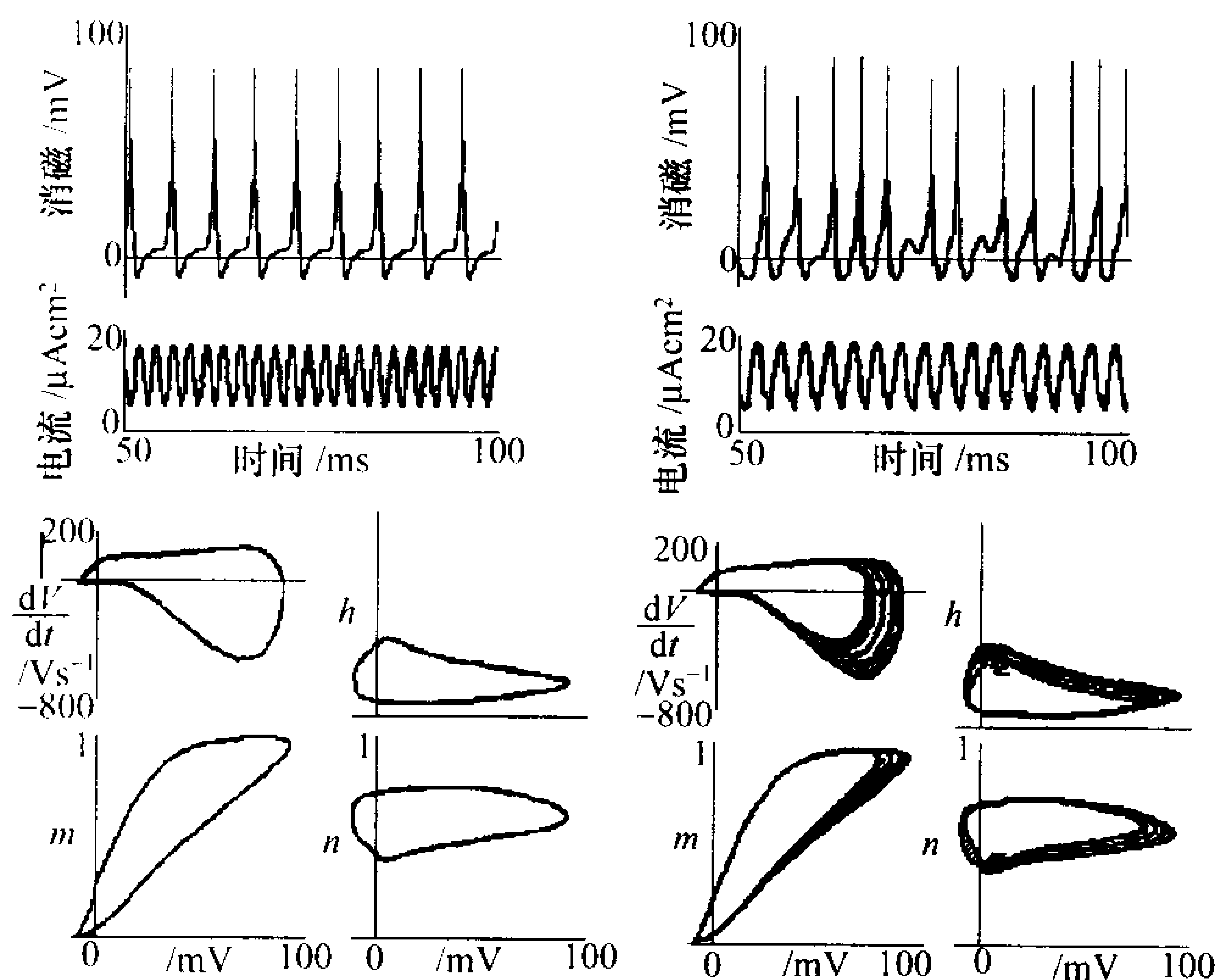
人们已得到图 3.6.19 中的无规律活动，该活动是周期驱动函数的确定微分系统的数值解，且不是随机的. 在一个实际试验中，同一个不规则活动可能是随机的，它产生于某些内外的噪音，或者是一个确定的混沌实例. 估计 Lyapunov 指数提供了一种识别混沌的方法.

有关鱿鱼轴突的试验给出了与 H–H 方程相同的行为，当用低 [Ca^{++}] 在轴突中建立一个自然频率时， H–H 方程可以作适当修正. 当一个频率为 itf_s 的正弦调节输入充分接近轴突的自然频率 f_n 时，响应可能变成了纳入；对于刺激电流的每个周期，轴突产生一个单作用电压. Aihara 等发现在一定电流和频率范围内响应



80
70
$V(t+T)$/mV
40
50
60
V/mV
(a)
45
40
35
30
$V(t+2T)$/mV
70
80
$V(t+T)$/mV
(b)
V
$V(t+T)$
$V(t+2T)$
(c)

图 3.6.18



100
消磁 /mV
0
20
电流 /μAcm^{-2}
0
50
时间 /ms
100
200
dV/dt
/Vs^{-1}
−800
h
m
n
1
0
/mV
100

图 3.6.19

仍保持纳入，并且也出现与 $\frac{f_s}{f_n}=\frac{m}{n}$ (这里 m 和 n 是整数) 保持同步. 大范围调节轴突迫使其远离其自然频率将激发混沌响应, 作用电压的振幅变得没有规律, 该行为与正弦迫动的 H–H 方程的数值解类似. 正弦迫动的鱿鱼轴突膜系统的最大 Lyapunov 指数是 97bits/s 或者 0.5bits/orbit.

3. 内生混沌

神经元的正弦迫动可产生混沌, 但是只能在迫动振幅与迫动频率较小的范围内, 因此没有实际的生物价值. 我们所感兴趣的是一个孤立的神经元能产生无规则的活动, 该活动被识别为混沌的. 有关软体动物神经元的试验表明通过钾离子通道阻塞剂在一个孤立的神经元上可以诱发混沌行为, 并且 Chay 方程的数值解的确表明了混沌的阵发. 然而这些只能在参数空间的一个很小范围内才发生.

§3.7 嗅觉神经网络

3.7.1 嗅觉系统的解剖学和生理学

嗅觉感受器位于鼻深部的粘膜之中, 这是一种双极细胞, 其远心端形成膨起, 称为 **终扣**, 是换能的主要所在. 终扣的末端上有很多纤毛, 埋于粘液之中. 吸气时, 嗅物质的分子扩散通过粘膜到达纤毛, 与膜上的受体蛋白结合, 开放内向正离子流通道, 从而产生去极化的感受器电位, 最终触发动作电位沿其近心端突起传导到达嗅球. 嗅觉感受器之间没有直接的相互联系, 它们的轴突构成初级嗅神经 (PON).

嗅球中的僧帽细胞是第二级神经元, 其主树突和初级嗅神经的末梢在一种称为小球的组织中形成突触联系. 小球之间通过小球周围细胞 (P) 相互联系起来. 小球周围细胞和 PON 形成轴树联系, 而其轴突则横向到达邻近的小球并与别的小球周围细胞或僧帽细胞形成轴树或轴体联系. 僧帽细胞接受来自初级嗅神经和周围小球细胞输出的信号, 彼此之间相互兴奋, 而其输出则通过侧嗅

束 (LOT) 分别到达前嗅核 (AON) 和梨状皮层 (PC) 的表层锥体细胞. 此外, 在嗅球深部颗粒层的颗粒细胞既接受来自僧帽细胞的兴奋性输入, 又反过来抑制僧帽细胞, 从而构成反馈回路而形成神经振荡器, 梨状皮层的输出最后通过深层锥体细胞进入外囊, 同时通过中嗅束 (MOT) 离中反馈回嗅球和前嗅核. 前嗅核也有反馈回颗粒层和周围小球层, 嗅觉神经网络的这种解剖结构如图 3.7.1 所示.

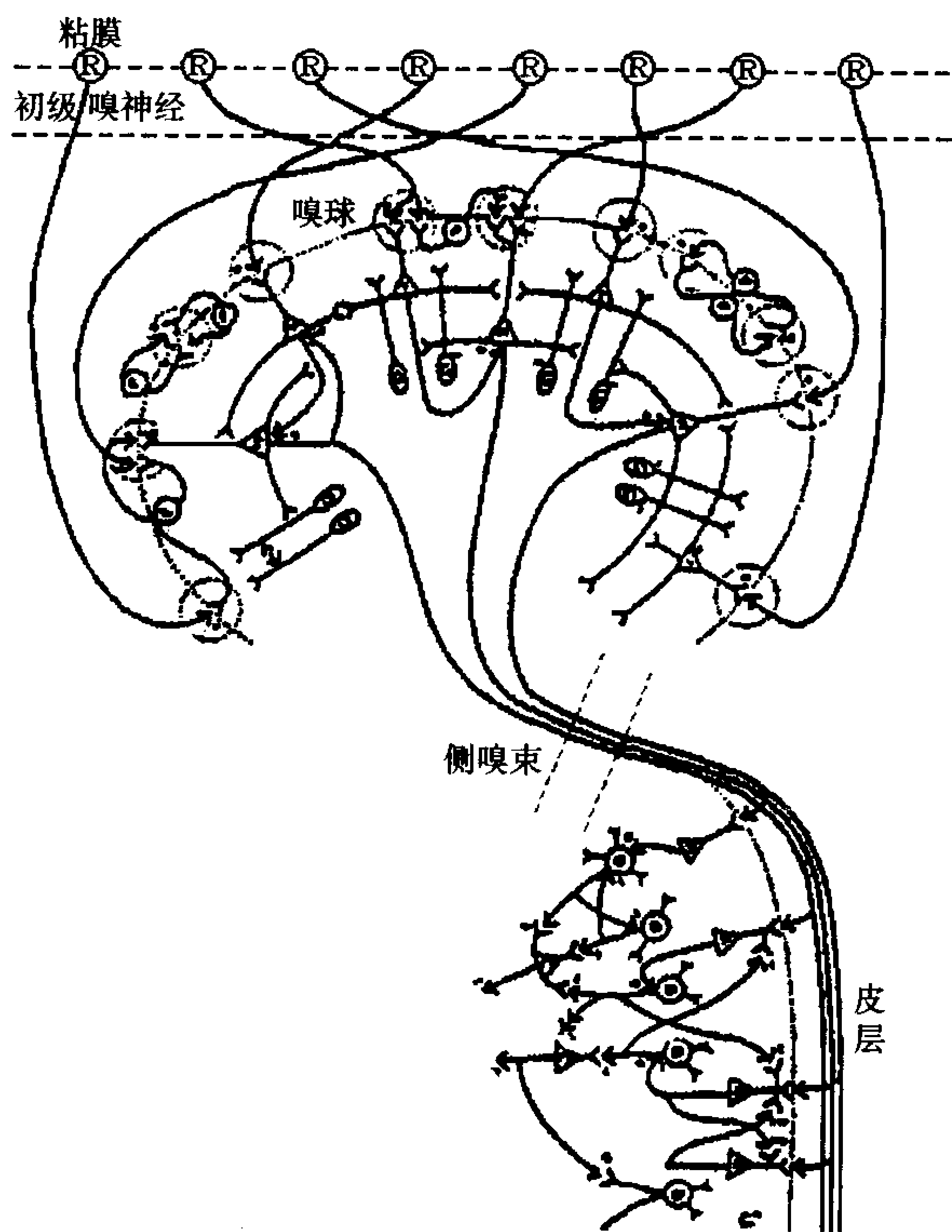


图 3.7.1 嗅球和梨状皮层中的主要细胞类型及其相互联系的示意图

(引自 Freeman, 1972)

现在已经知道, 一个嗅感受器可以对许多不同的嗅味产生反应, 只是敏感度可能不同. 而同一种嗅味只能引起许多感受器的反应, 却没有发现各级中枢的神经元中有比感受器对某种嗅味的调

谐特性更好的. 此外, 至今也不知道是否存在有最基本的嗅味, 因此嗅觉感知并不取决于单个神经元的活动. Freeman 通过多年的研究发现, 决定嗅觉感知的是大量神经元的集体活动, 在嗅脑电的空间分布模式中携带有这种信息.

他们的实验是这样进行的: 对口渴的兔子进行条件反射实验, 以某种嗅味作为条件刺激 (CS+), 隔 2 秒后给兔子水喝, 而以兔子的舔舌动作作为条件反射 (CR+); 以另一种嗅味作为刺激而不给水喝 (CS–), 而以嗅味的动作作为反应 (CR–). 在兔子左嗅球的外侧表面上埋藏 8×8 个电极阵列, 将由此引导得的 64 道脑电放大和滤波. 当兔子嗅进一种熟悉的气味时, 各道脑电的波形都突然变得更为规整和有序, 直到呼气为止, 而且其幅度和频率都较他时为高, 我们称之为 **阵发性波动** (burst) (见图 3.7.2). 此外观察吸气时的波形还可发现, 各道记录的波形彼此相似, 但幅度各不相同 (见图 3.7.3). 以各道记录中基本波形的平均值作为指标画出地形图, 就可得出某种空间分布模式. 每个动物在对某种嗅味熟悉以后, 尽管每次吸气时这种载波波形都不一样, 但其幅度空间模式却是相当恒定的. 对于不同的嗅味, 相应的空间模式也不同, 看来正是在这种空间分布模式中携带有有关嗅味特性的信息. 当给动物嗅一种新的气味 CS+, 或将原来的气味 CS+ 改变为 CS–, 整个空间分布模式才会改变 (见图 3.7.4). 因此这种空间模式不仅和刺激有关, 而且还和动物以前的经验有关, 赋予刺激以某种意义, 亦即和经验有关.

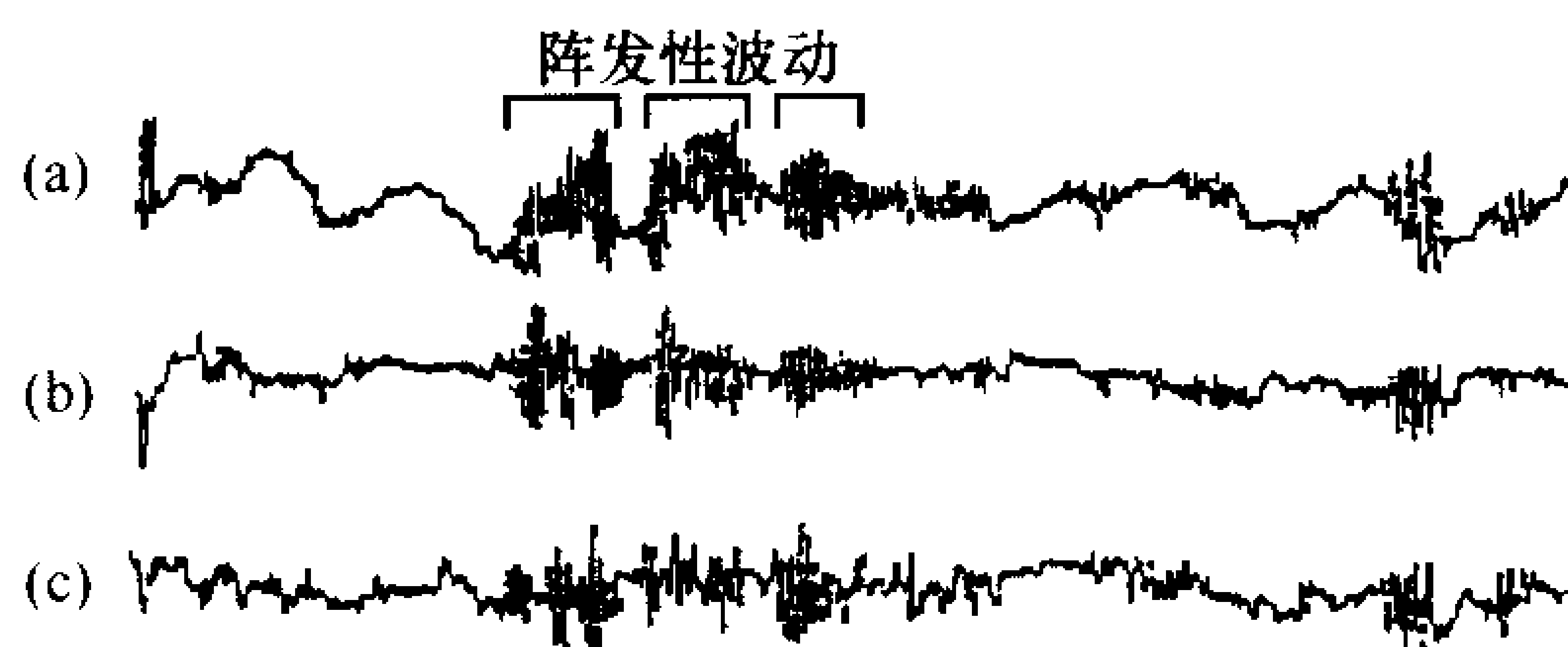


图 3.7.2　由猫嗅球 (a), 前嗅皮层 (b), 后嗅皮层 (c) 同时记录到的脑电, 注意当猫吸进它知道的气味时, 原来的低幅慢波变成高幅, 高频的振荡

(引自 Freeman, 1991)

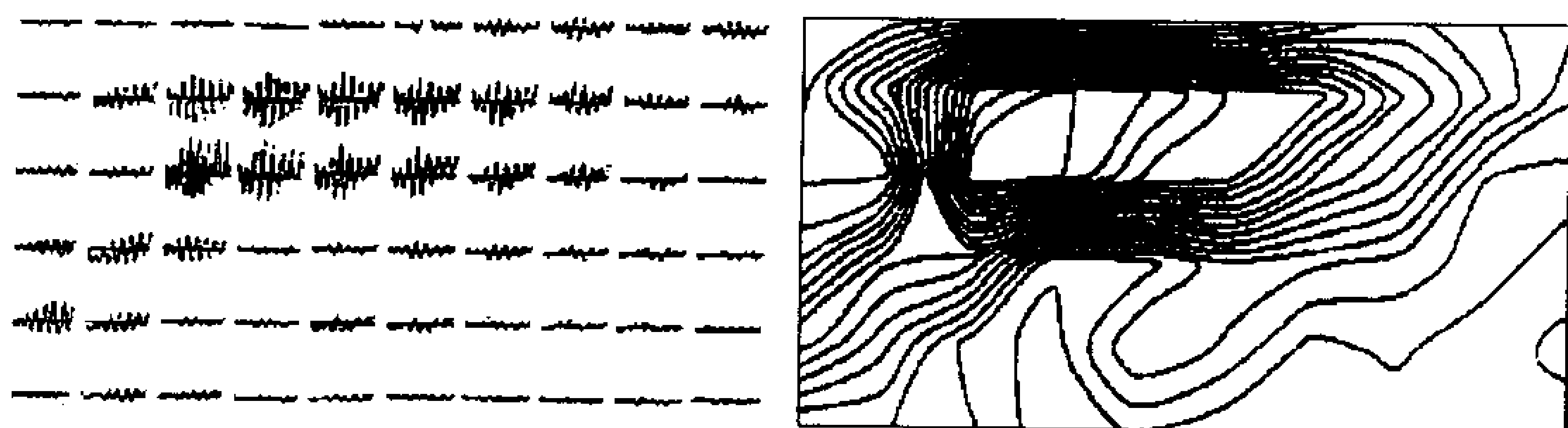


图 3.7.3　左图给出当兔子辨认出一种气味时，从嗅皮层上同时记录到的 60 道脑电. 由图可见这些脑电的波形几乎相同，但幅度各不相同. 右图是按这些波形的幅度画出的地形图. 它携带有嗅入气味特性的信息 (引自 Freeman, 1991)

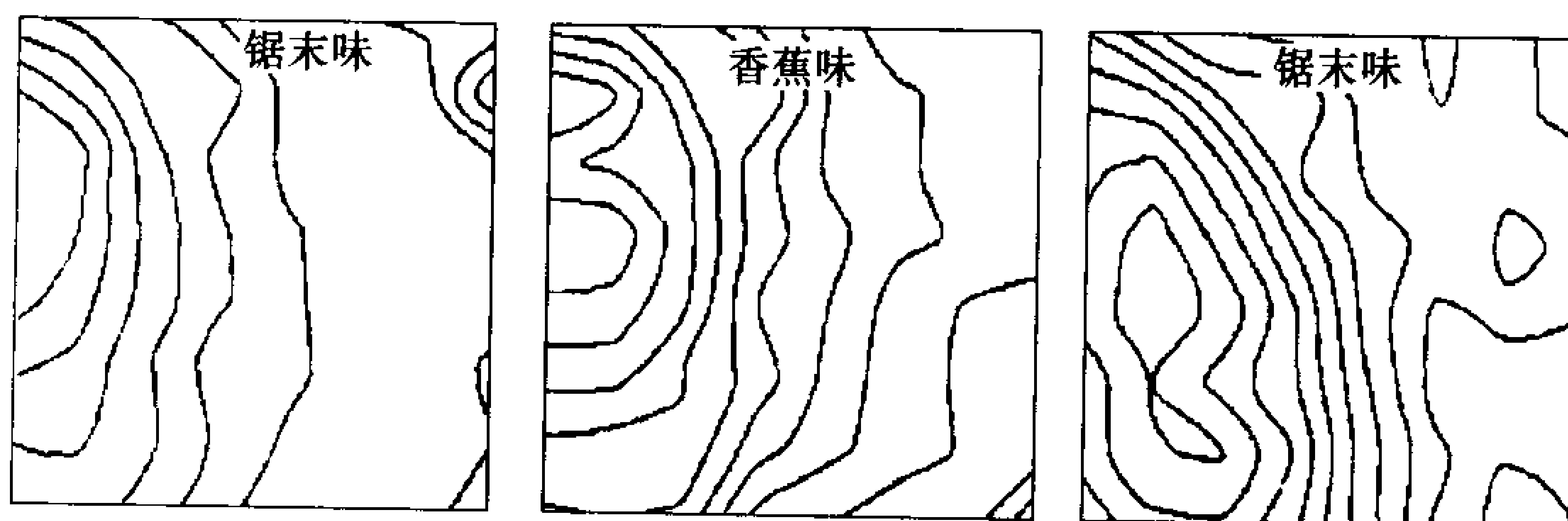


图 3.7.4　左图为兔子对锯末味建立条件反射以后由嗅球上记录到的脑电图的地形图，然后让兔子学习辨识香蕉味 (中图)，此后再让兔子嗅锯末味，其地形图发生了很大的变化 (右图)，这说明经验对嗅球活动的影响比刺激更大，否则地形图将维持不变 (引自 Freeman, 1991)

3.7.2　嗅觉神经网络模型

根据嗅觉系统的解剖结构， Yao 和 Freeman (1990) 给出了嗅觉神经网络的模型框图 (图 3.7.5). 图顶上的箭头来自感受器，这是系统的输入. 箭头边上的 + 号表示兴奋性， − 号表示抑制性，L_1–L_4 表示延误，每个圆圈表示一群神经元. 生理实验表明，每个神经元群可以用一个输出经过 S 形函数作用的二阶线性微分方程来表示. 每条通道都赋予一个增益值 K，其足标表示通道的起点和终点. 例如， K_{pr}, K_{mr} 分别表示从 R 到 P 和从 R 到 M 的增益值. $K_{mm}[i,j]$ 是联结从第 j 个单元中的 M 细胞到第 i 个单元中的 M 细胞的增益. K_{gglink} 是联结两个单元的 G 细胞之间的联结权重. 双箭头表示来自该水平所有神经元的信号或是送到该水平所有神经元去的信号.

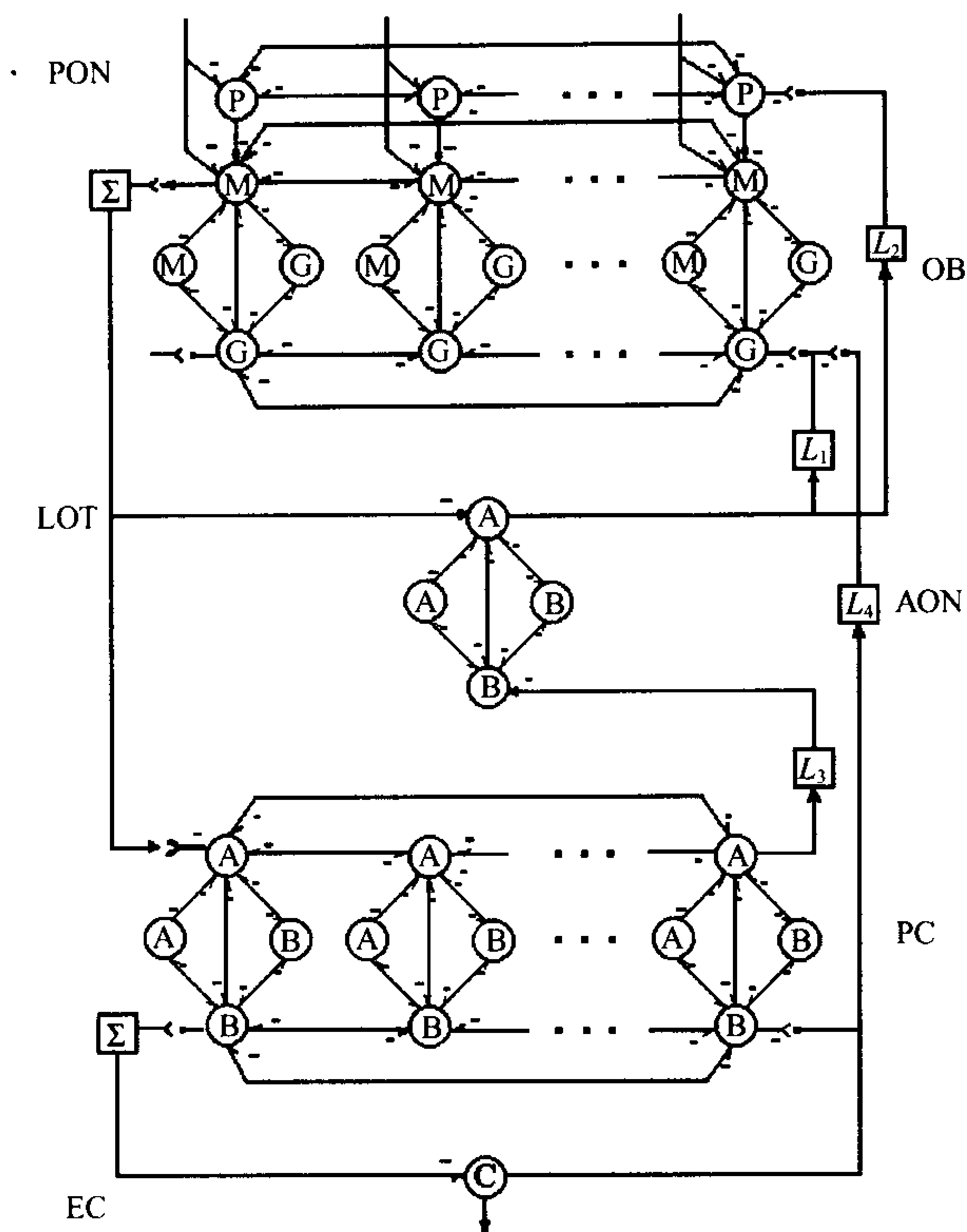


图 3.7.5 嗅觉神经网络的模型框图

(引自 Freeman, 1990)

符号: P —— 小球周围细胞;

M —— 僧帽细胞;

G —— 颗粒细胞;

OB —— 嗅球;

AON —— 前嗅核;

PC —— 梨状皮层.

在此模型的联系中, 只有兴奋性神经元 (M 或 A) 之间的兴奋性联结才是用于学习的可修改权重, 学习采用相关算法.

由上述可知，此模型在解剖学上是真实的 (见图 3.7.6). 对模型进行训练使它认出某种气味之后，当给模型加上该种输入时，嗅球部分就会产生阵发性波动 (见图 3.7.7)，而且其公共载波形成明显而恒定的幅值地形图. 让模型学会认出新的气味之后，除了产生新的地形图之外，原来的地形图也会随之而变.

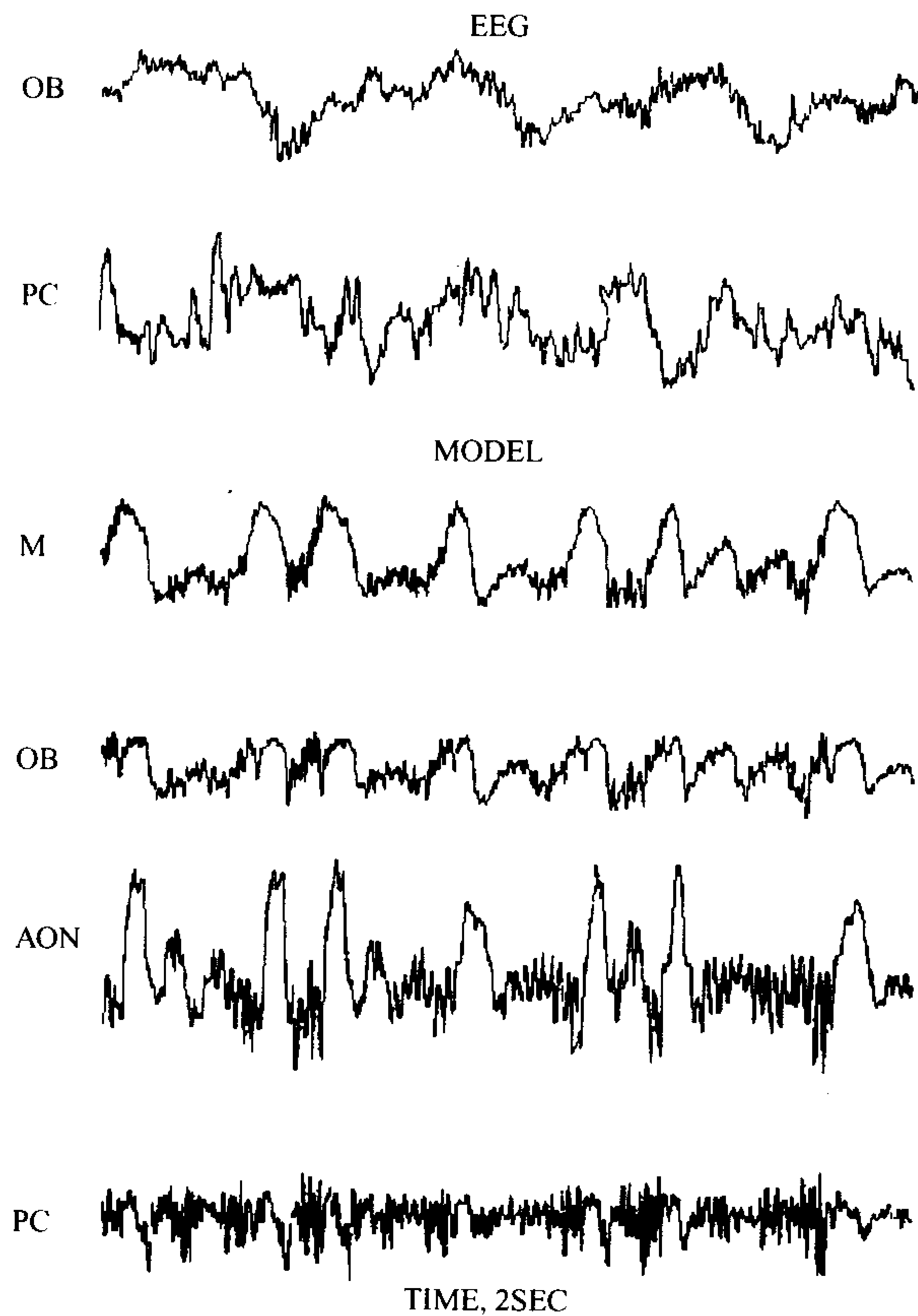


图 3.7.6　在嗅球 (OB) 和皮层 (PC) 上记录到的基础活动 (EEG, 上图) 和模型产生的非周期背景活动 (MODEL, 下图). 下图中曲面由上而下分别仿真嗅球单元活动 (M)，以及在嗅球 (OB)、前嗅核 (AON) 和梨状皮层 (PC) 的脑电

(引自 Freeman, 1987)

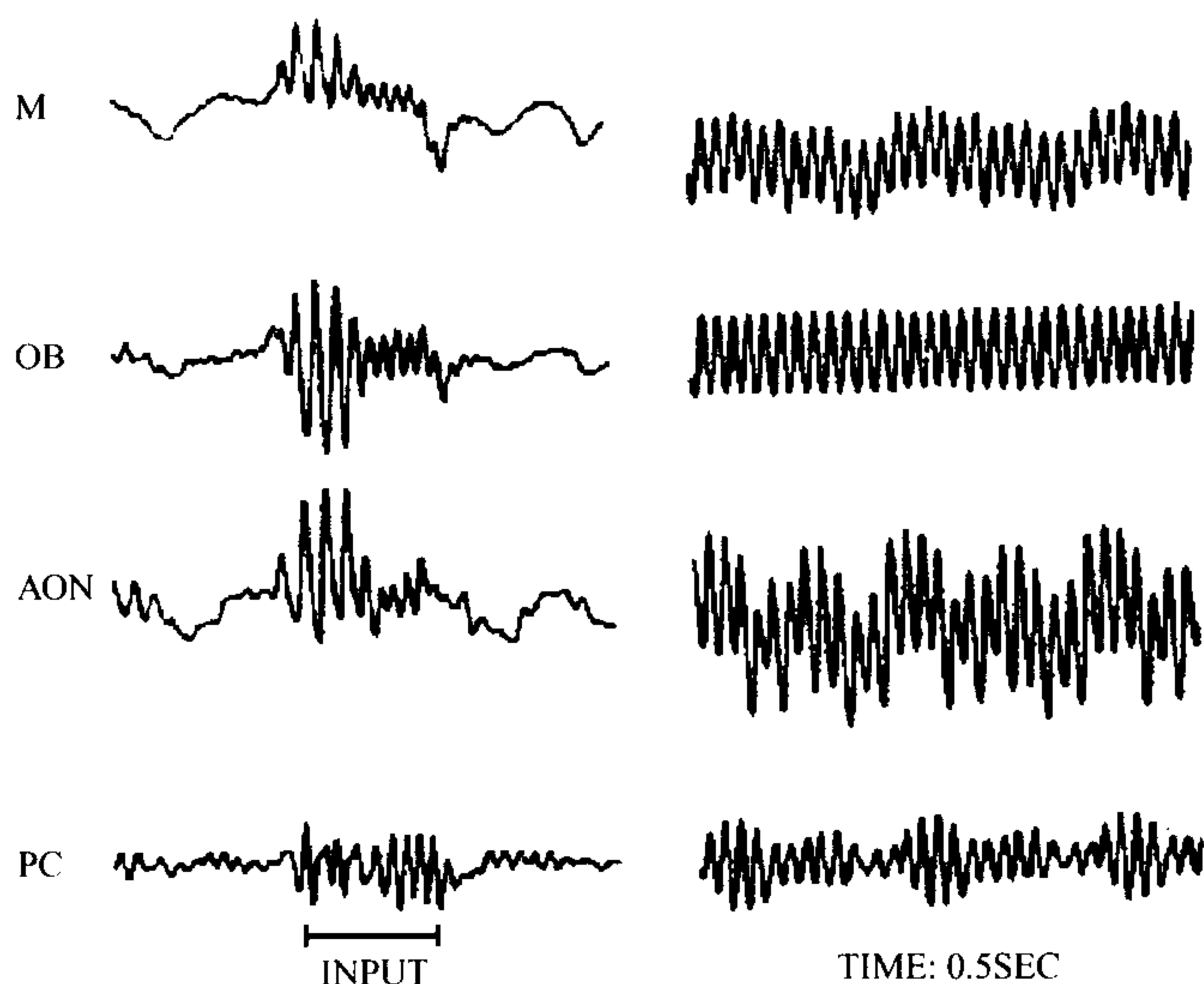


图 3.7.7　模型上 R 处给予一段输入 (吸气 0.2s) 所诱发的仿真阵发性波动 (左图). 右图为持续输入下的活动

(引自 Freeman, 1987)

3.7.3　嗅觉系统中存在混沌的证据

嗅觉系统中存在混沌的第一个迹象是即使没有外部驱动力, 嗅球各处也总是存在非周期的公共载波. 只有用手术切断嗅球和梨状皮层以及它们和感受器及脑的其余部分的联系, 或是深度麻醉或冷冻阻断嗅球和梨状皮层之间的联系才能使它们的活动归于平静, 亦即处于稳定的平衡态. 然而只要在嗅球和梨状皮层之间保持完整的联系, 它们就不能停留在平衡态, 而是进入无休止的活动之中. 用 Grassberger–Procaccia 算法 (1983) 计算这种脑电的 Hausdorff 分维数所得的值在 4 到 7 之间 (Freeman, 1987). 全身麻醉可以阻抑这种活动, 清醒之后又回复这种活动. 给予瞬间的电刺激可以干扰这种活动, 然而很快又恢复原来的状态. 看来这种状态是一种稳定的混沌吸引子. 这种活动被称之为 **基础活动** 或 **自发活动**.

第二个迹象是嗅球和皮层活动可以非常快地整个从一种状态跳变到另一种状态. 这种状态转换既可以在没有输入时发生, 例如动物由清醒状态转入睡眠状态, 也可以通过给予输入产生. 如前所述, 当动物吸入一种熟悉的气味时, 脑电就从基础活动转入阵发性波动, 这表明此时嗅觉系统失稳, 它离开了它的基础混沌态 ——一种宽带混沌, 而转入一种窄带混沌.

这种状态变化的一个病理例子是癫痫. 虽然癫痫有许多不同的类型, 其复杂程度也各不相同, 但它们的公共特征是: (i) 突然地而且全局地产生宽频带、高幅度的非周期活动; (ii) 丧失对刺激的反应性, 也丧失有目的的行为; (iii) 突然地和全局地终止这种状态. 对动物可以用短暂而且高强度的电刺激从 OB 到 AON 和 PC 的传导通路来诱发癫痫. 在 10 ~ 100 脉冲 /s 刺激几秒钟之后, 脑电活动突然降低到很低的水平, 这表明系统接近它的一个平衡点. 经过几秒钟之后, 以约 3 次 /s 出现脉动波 (见图 3.7.8), 这表明这个平衡点是不稳定的. 虽然这种脉动波看来似乎是周期的, 但是如果把由 OB 和 AON 上记录到的脑电画成相图, 那么其轨线就不是极限环, 而是一种混沌吸引子.

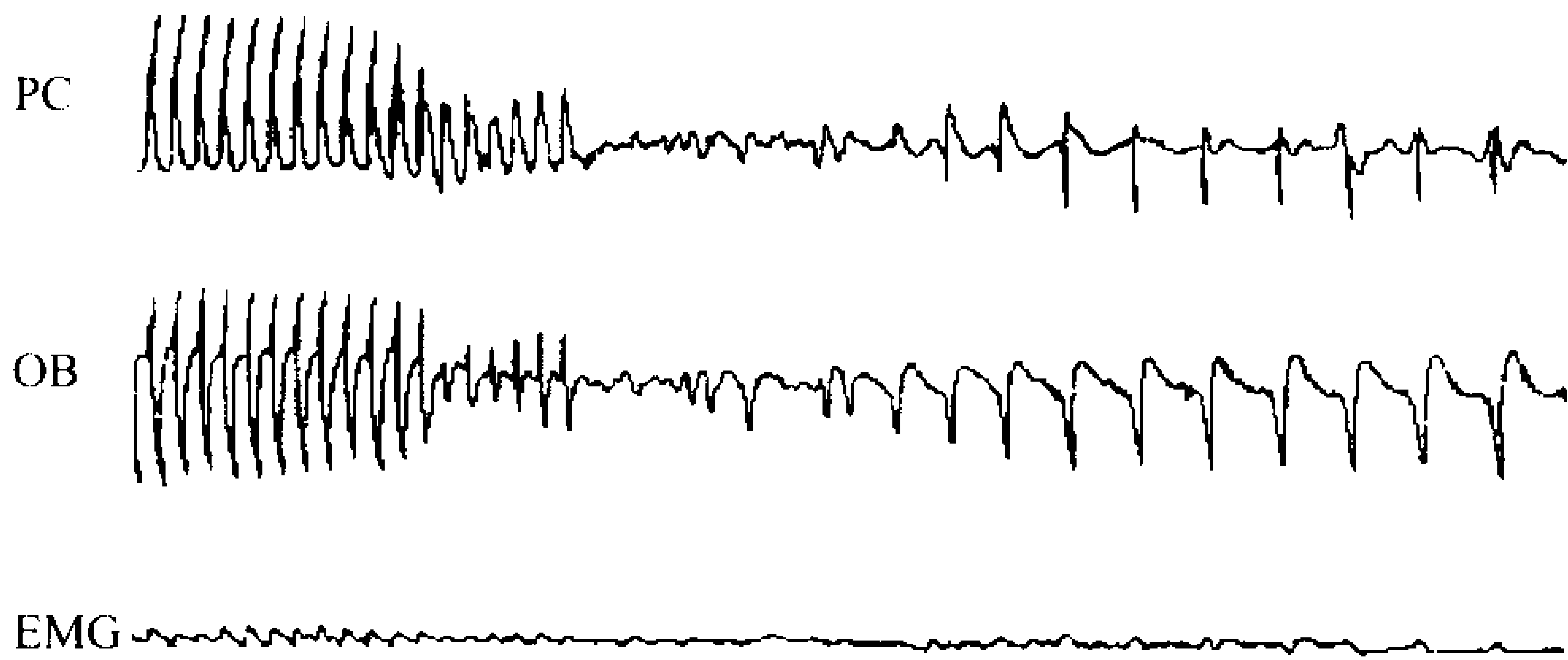


图 3.7.8　癫痫发作前后由嗅球和皮层上记录到的脑电以及由下颌处记录到的肌电 (EMG)

(引自 Freeman, 1987)

上一节中已经说明, Freeman 等人所建立的神经网络模型无论从解剖上或是从生理上都有其真实性, 因此研究模型的非线性动力学性质也是论证嗅觉系统的混沌性质的一种途径.

在模型中将联结从 OB 到 PC 的增益减少到正常值的 15%, 就

可仿真癫痫活动 (图 3.7.9).

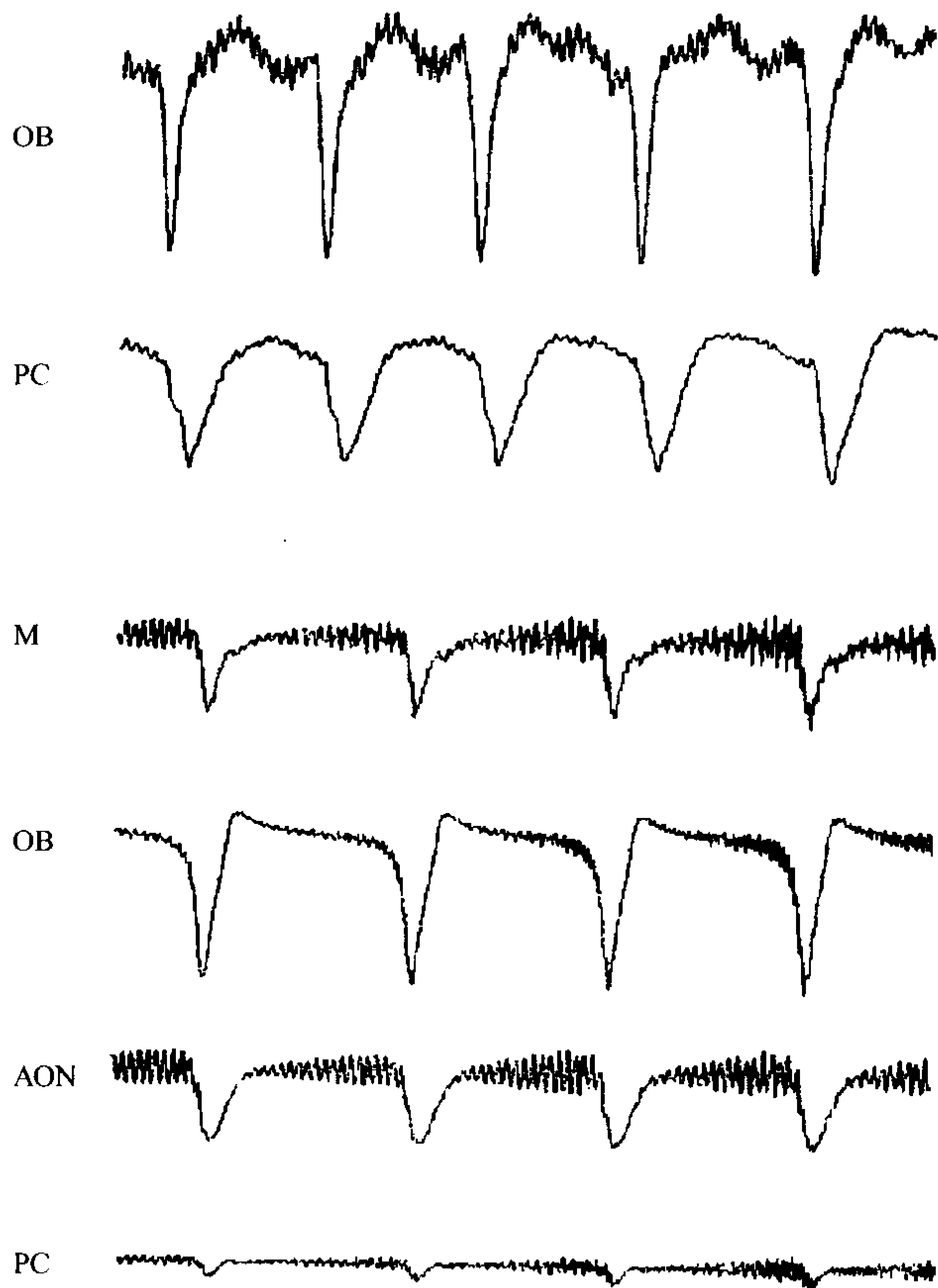


图 3.7.9 癫痫发作时的脑电 (EEG) 及其模型仿真 (MODEL)

(引自 Freeman, 1987)

Freeman (1987) 发现模型中联结互兴奋性子集合 (P=P) 和互兴奋性子集合 (M=M) 的增益是一个关键参数. 当 $K_{pm}=0$ 时整个系统趋于稳定平衡态, 这一状态仿真了处于深度麻醉时, 3 个部分 "开环" 的状态. K_{pm} 增加到 0.1, 系统经过一个 Hopf 分岔而达到约为 3 次 /s 的低频极限环. 当 K_{pm} 增加到 0.2 时, 经过

第二个 Hopf 分岔，极限环处于环面的表面上. 当 K_{pm} 增加到 0.3 时，环面崩塌而出现混沌. 将 OB 到 PC 的增益从 0.15 增大到 1.0 就可使其活动进一步分岔到基础混沌吸引子.

计算实验测得和模型仿真所得的癫痫脑电相关维数，其值分别为 2.52 和 3.76, 相应的基础脑电相关维数分别为 5.46 和 5.92, 相应的阵发波动相关维数值分别为 4.10 ~ 4.88 和 3.77, 这些发现表明癫痫发作状态和阵发波动状态的复杂性要比基础状态低.

以模型中的一些单元作坐标画相图，可以看出其轨线有明显的结构，这提示了系统的混沌性质 (见图 3.7.10).

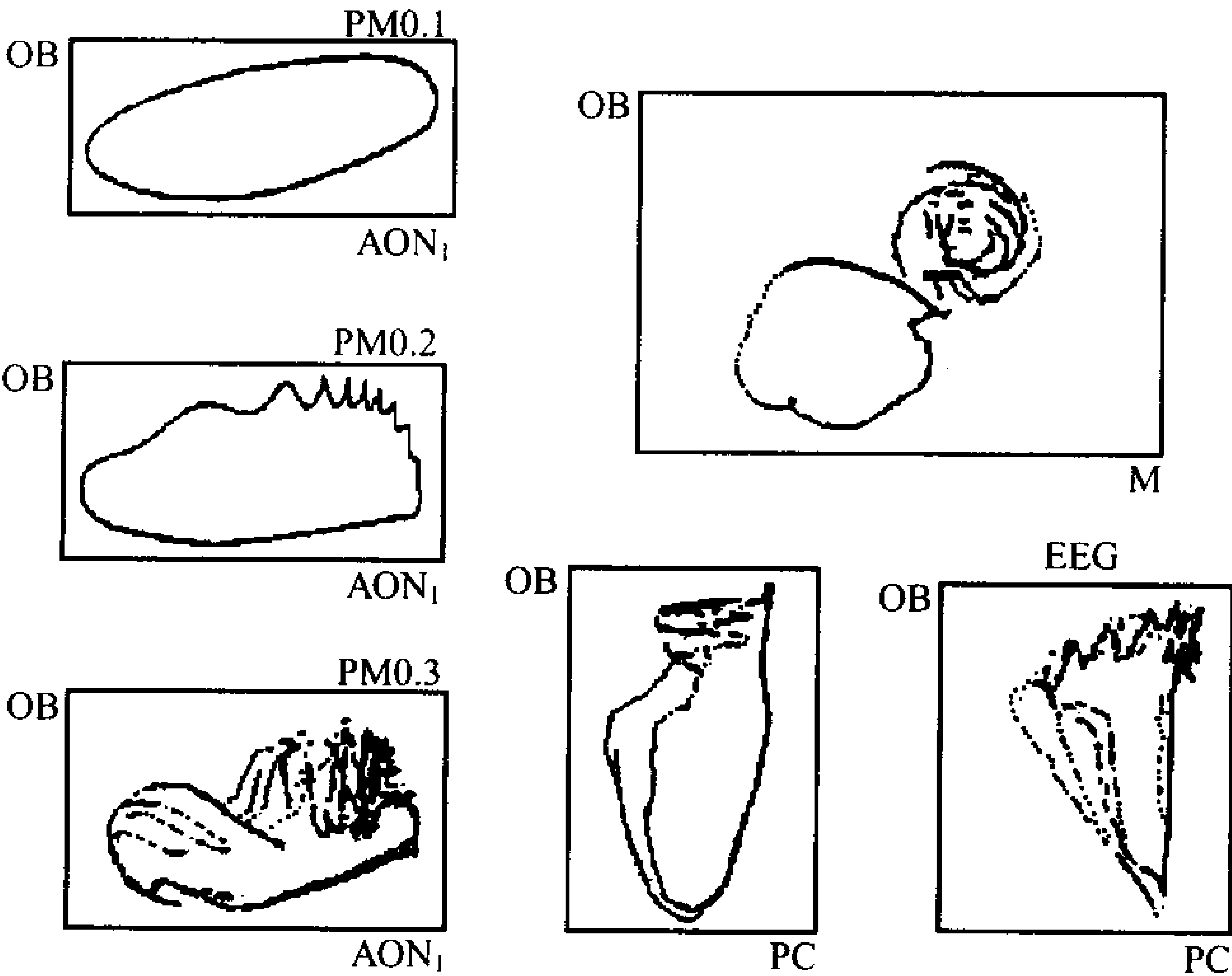


图 3.7.10 嗅觉神经网络模型的相图

(引自 Freeman, 1991)

3.7.4 混沌在嗅觉信息处理中的可能作用

模式识别是神经系统擅长而计算机实现困难的一种功能. 在传统的人工智能中，模式辨认是通过特征提取、摸索和模板匹配来实现的，这种方法要消耗大量机时. 在生理学中关于感觉信息处理的传统观念认为，对象的各种特征由某些称为特征检测器的神经元加以检测，最后由某个称为“主教细胞”或“老祖母细胞”的神

经元加以辨认. 这种理论缺乏大量的实验证据, 而且在理论上遇到组合爆炸的困难. Freeman 的研究提出了一种新的途径: 动物所能辨认出的每一种气味都相应于混沌吸引子的某些翼, 而其静息状态相当于混沌吸引子的核. 动物认出某种气味就是跳到相应的翼. 当动物学会辨认一种新的气味, 就产生一个新的翼, 而原来的翼也要略做改变.

目前的研究还不能证明嗅觉系统中确实存在混沌, 也不能确切地分清混沌现象究竟对其功能是否必需, 还仅仅只是一种副产品. 但无论如何上述工作对感觉信息处理提出了一种新的想法, 这种想法克服了传统理论的一些本质困难, 无论对工程技术还是神经学都开辟了新的途径.

§3.8 生物神经网络中的同步振荡

3.8.1 引言

生物神经网络中的同步振荡问题实质上是一个非线性振子之间的耦合和相位锁定问题. 由于它对理解感知的重要意义, 因此也引起非线性科学家的重视.

我们怎样感知周围世界中的对象, 这是感觉生理学的中心问题, 也是至今尚未解决的一个难题. 视觉生理学已有大量的实验事实表明视觉器官有可能首先抽提对象的各个局部特征并且并行地向中枢传送, 问题是中枢怎样再将这些特征重新整合起来, 这就是所谓的 **结合问题** (binding problem). 对象和背景的分割以及同一背景中不同对象的区分都属于这类问题. 一开始人们提出了一种 **老祖母细胞** (或称 **主教细胞**) 的理论, 这种理论认为检测不同特征的神经元的输出在向中枢传送的过程中分层次地收敛于越来越异化的神经元, 最终每个对象都有一个神经元对之起反应. 但是这种理论遇到所谓组合爆炸的困难, 即为了对周围世界中每一种可能的对象起反应所需要的神经元数目远远超过大脑所能容纳的数量. Hebb 提出的另一种理论可以解决这一困难. 他认为为

了检测一个对象，需要有一群神经元提高它们的平均发放量. 这一理论虽然克服了组合爆炸的困难，却又引起新的问题，即它无法区分同时呈现的不同对象，这个问题被称为**叠加灾难** (superposotion catastrophe). 为了克服这一困难， von der Malsburg 和 Schneider (1986) 对 Hebb 的理论进行了修正. 他们的理论要点如下：同一对象由一群同步振荡的神经元来表征. 属于同一群体的神经元的发放彼此同步，而不同群体的神经元的发放则不存在一致的相位关系，因此利用相位信息可以区分不同的对象，并有选择地把属于同一对象的分步在各处的特征结合在一起. 因此他们的模型像 Hebb 模型一样避免了组合爆炸的困难，而且由于只有发放的时间模式彼此同步的神经元才属于同一群体，因此它又克服了叠加灾难. 此外由于同一神经元在不同的时间可以和不同的神经元同步，因此这种表达方法十分经济.

3.8.2 生理学证据

von der Malsburg 的理论在开始时只是一种假设，对这一假设的第一批实验支持来自 Freeman 关于兔嗅球和嗅皮层的实验. 他们的实验表明由兔和猫的嗅球和嗅皮层确实可记录到频率在 $35 \sim 90$Hz 之间的振荡电位. 以后 Singer 和 Eckhorn 两个实验室又分别在猫视皮层上记录到频率在 $40 \sim 60$Hz 之间的局域场电位，单单元活动和多单元活动. 这种振荡性质可通过对数据进行谱分析或自相关分析而得到证实.

他们的进一步研究又表明对同一对象起反应的各个神经元的活动是同步的，这种同步性可以通过计算它们之间的相关函数加以检验. 这种同步振荡的现象不仅可以在同一功能柱内的细胞之间记录到，而且在同一皮层区内的相隔达 7mm 的细胞之间、同一半球的不同皮层区之间、或不同半球的对应部位之间都可以记录到.

特别有意思的是不同细胞活动之间的不同步或同步取决于引起它们反应的刺激之间的相干程度，图 3.8.1 就是一个这样的实验结果. 从猫视皮层 17 区彼此相隔 0.4mm 的 4 个有不同朝向的功

能柱中分别记录多单元活动. 这 4 个细胞群的感受野彼此重叠, 且其最优朝向分别为 22 度 (群 1) 、 112 度 (群 2) 、 157 度 (群 3) 和 90 度 (群 4), 这些朝向分别用图 3.8.1 的 (a),(b),(c),(d) 中穿过每个感受野中心的粗线来表示. 左图表示用各种不同朝向的单个光条来回移动时的反应, 右图表示有两个不同朝向的光条同时运动时的反应. 在每一种刺激条件下, 有反应的细胞群用填满阴影线的感受野来表示. 单光条刺激只能引起有相近朝向的神经元的同步活动. 因此, 图 (a) 中只有群 1 和 3 对垂直朝向 (0 度) 的光条同

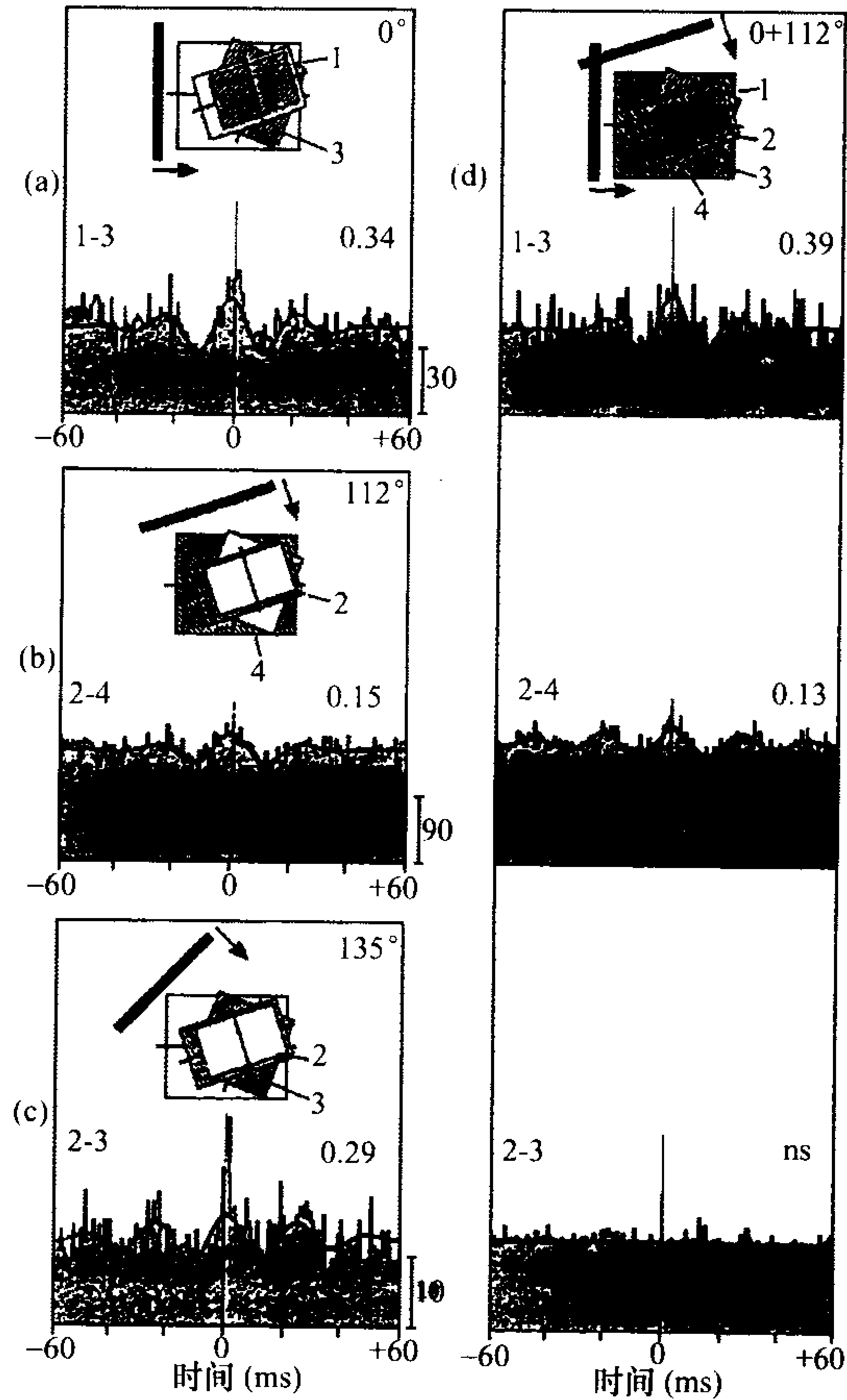


图 3.8.1　有重叠感受野的神经元群对不同刺激的反应 (说明详见正文)

步活动; 图 (b) 中只有群 2 和 4 对朝向为 112 度的光条同步; 图 (c) 中只有群 2 和 3 对朝向为 135 度的光条同步. 当同时在感受野中有两个朝向分别为 0 度和 112 度的光条同时运动时, 所有 4 群细胞都被激活 (图 (d)). 然而这 4 群细胞按照其最优朝向被分为两组: 群 1 和群 3 的最优朝向接近于垂直刺激光条, 它们的活动彼此同步而构成一组; 群 2 和群 4 的最优朝向接近于 112 度, 它们的活动也彼此同步而构成另一组. 但是这两组彼此之间互不同步. 例如群 2 和群 3 的互相关图就几乎是一条水平线. 图 3.8.1 是另一个实验结果. 图 3.8.1(a) 表示由相隔 7mm 的两个最优朝向为垂直方向的细胞群 (1,2) 处测得的朝向调谐曲线. 图 3.8.1(b) 是在 3 种实验条件下在该两处分别测得的刺激后时间直方图: (i) 两个光条反向运动; (ii) 两个光条同向运动; (iii) 一根长光条同时通过这两个感受野运动. 感受野的位置和刺激条件都图示于每张刺激后时间直方图的右边. 图 3.8.1(c) 是该两处在 3 种实验条件下测得的自相关图, 而图 3.8.1(d) 则是在这 3 种条件下测得的该两处的互相关图.

图 3.8.2 是当感受野互不重叠时的神经元群对 3 种不同刺激形式的反应. 这两组神经元群的最优朝向都是垂直方向 (0 度或 180 度), 其朝向调谐曲线如图中 (a) 所示. 图 (b) 表示这两组神经元群在 3 种不同刺激形式之下的刺激后时间直方图. (i) 是当两个垂直光条反向运动越过这两个感受野时所得出的刺激后时间直方图; (ii) 是当两个垂直光条同向越过这两个感受野时的反应; (iii) 则是当这两个垂直光条相互连接形成单独一个长光条时同时越过这两个感受野时的反应. 刺激形式图示在相应的刺激后时间直方图右侧. 图 (c) 是这两组神经元群分别在这 3 种情形下的自相关图, 而图 (d) 则表示这两组神经元群分别在这 3 种情形下的互相关图. 图中填黑和空白表示运动方向相反时分别得到的结果. 由图 (c) 自相关图中都有周期峰的出现可知这两个神经元群在所有 3 种情形下都表现出振荡行为. 而在图 (d) 的互相关图中可明显看出仅当两个刺激光条连成一片 (iii), 亦即只有一个刺激整体时, 神经元群 1 和神经元群 2 的活动才表现出强烈的同步行为. 当两个垂直光条

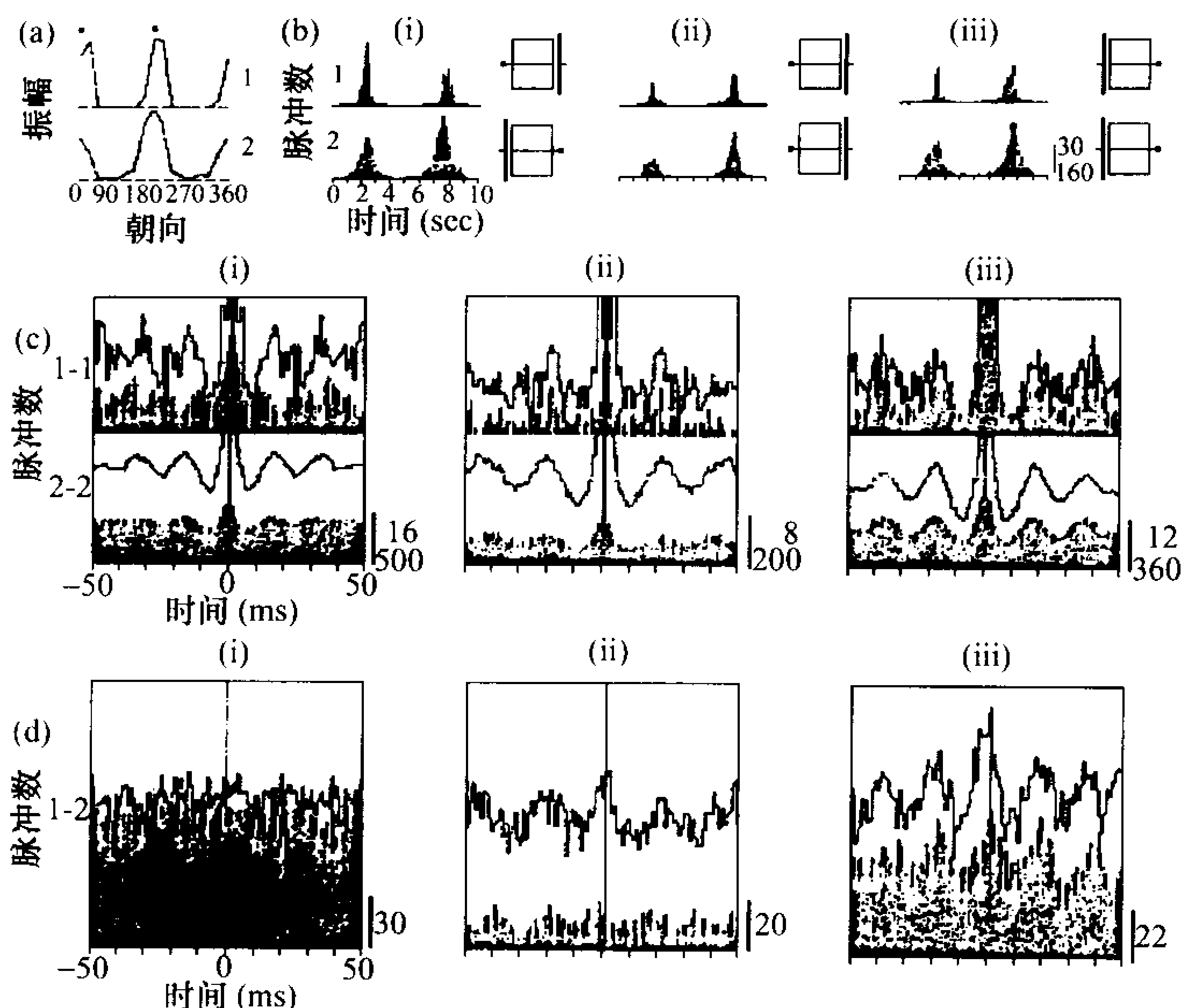


图 3.8.2　感受野互不重叠的神经元群对不同刺激的反应 (说明详见正文)

(引自 Grayetal, 1989)

分开同向运动时, 互相关减弱, 但仍表现出一定的同步性. 而当两个分隔开的光条反向运动时, 亦即存在两个不同的刺激分别刺激此二神经元群时, 它们的活动就不再彼此同步了.

3.8.3　神经网络仿真

Schillen 等人用一个带有时滞的非线性振子相互耦合构成的神经网络来仿真上述现象. 这种基本的非线性振子的结构如图 3.8.3 所示. 它是由一个兴奋性单元和一个抑制性单元相互通过有时滞的联系联结起来, 而在兴奋性单元上加上外界刺激. 在这里每个单元都代表一群神经元, 而其输出则为联合发放概率. 这个振子的动力学可以用下列带有时滞的非线性微分方程来描述:

$$\tau_0 X_e(t) = -\alpha_e X_e(t) - W_{ie} F(X_i(t - \tau_{ie})) + I_e(t) + \eta(t),$$

$$\tau_0 X_i(t) = -\alpha_i X_i(t) - W_{ei} F(X_e(t - \tau_{ei})) + \eta(t),$$

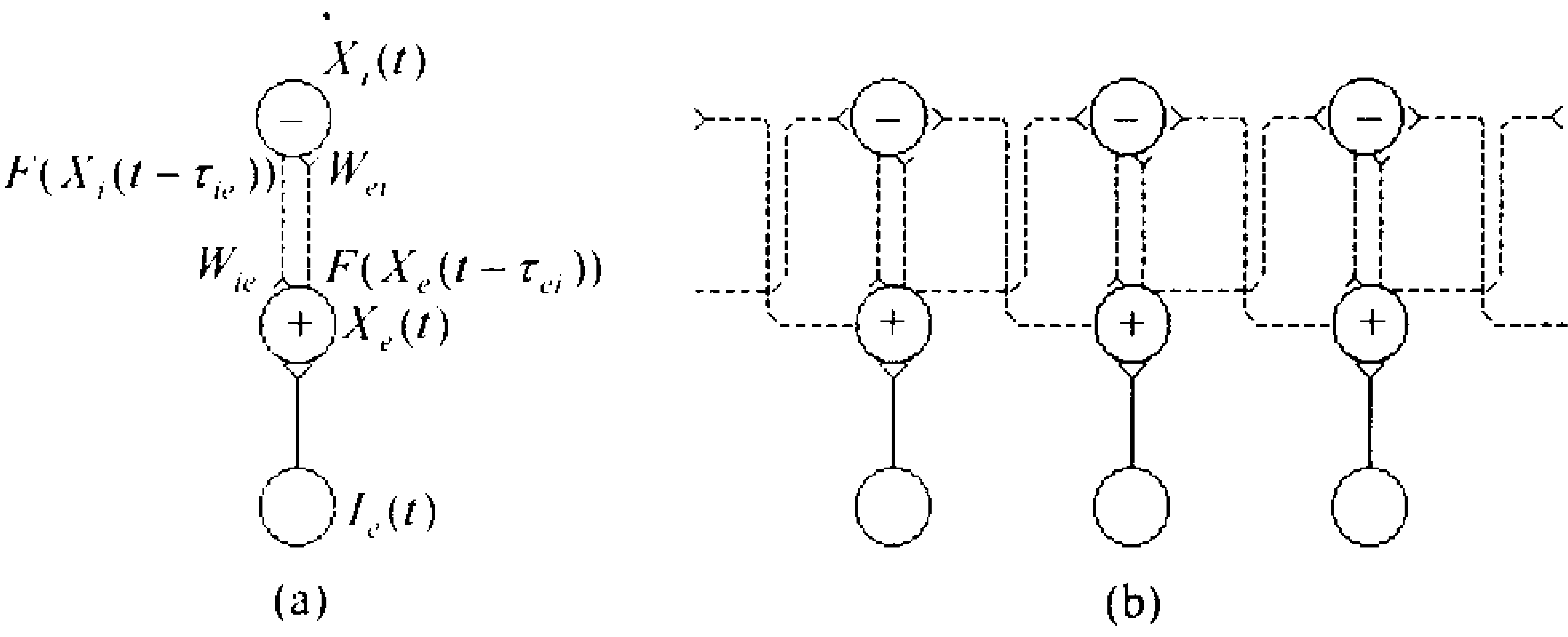


图 3.8.3

(a) 基本非线性振子的框图 (说明详见正文, 引自 Schillen, 1990); (b) 基本振子之间的同步性联结方式. + 号表示兴奋性神经元群, − 号表示抑制性神经元群

其中 t 为时间, $X(t)$ 为单元的活动量, α 为阻尼系数, W 为耦合系数 $(W > 0)$, τ 为时滞, $I_0(t)$ 为外界刺激, $F(X(t))$ 为 S 形函数

$$F(X(t)) = \frac{1}{\exp(\sigma(\Theta - X(t)) + 1)},$$

$\eta(t)$ 为白噪声.

Schillen 等通过将相邻振子的兴奋性单元和抑制性单元以带有时滞的连线联结起来, 实现了网络各个单元活动的同步化. 图 3.8.4 是仿真结果, 画出了 14×7 个单元通过上述联结构成的神经网络

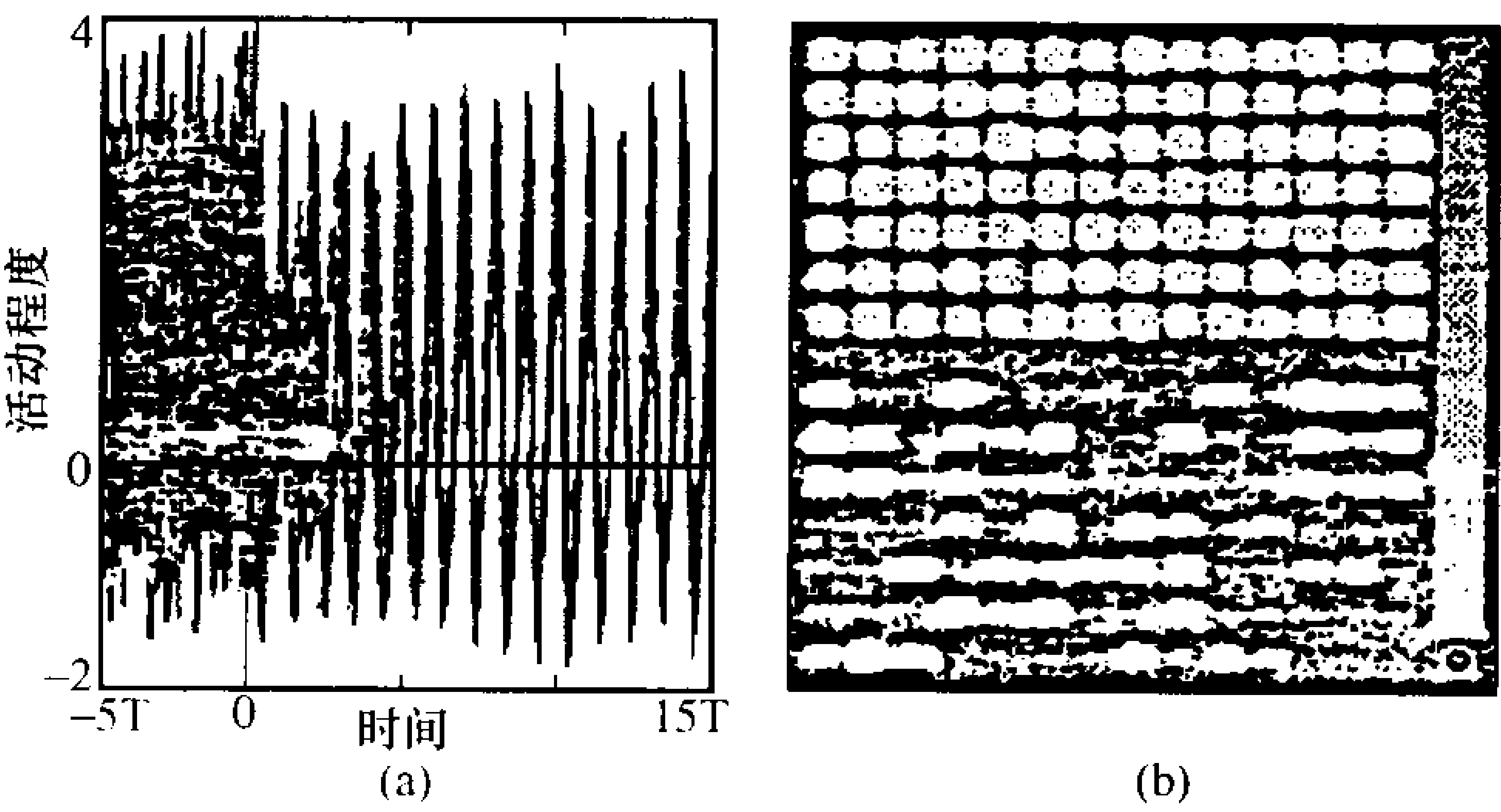


图 3.8.4 网络活动的同步化 (说明详见正文)

(引自 Schillen, 1990)

中任选 20 个兴奋性单元的活动. 当 $t < 0$ 时, 将网络中的所有联结断开, 并加上很强的噪声, 此时各个单元的活动互不同步. 当 $t > 0$ 时, 保持噪声强度不变, 但将相邻兴奋性单元之间用有时滞的连线联结起来, 可以看到, 各单元同步活动.

将振子的兴奋性单元之间用有时滞的连线联结起来, 可实现去同步化 (见图 3.8.5). 图 3.8.6 表示由 14×7 个振子中任选 20 个兴奋性单元的活动. 当 $t < 0$ 时, 振子之间彼此孤立, 但令它们的初始条件相图以保证其活动同步, 并加上小的噪声以打破对称. 当 $t > 0$ 时去除噪声, 然后把各兴奋性单元之间的联结加上, 此时即实现了去同步化.

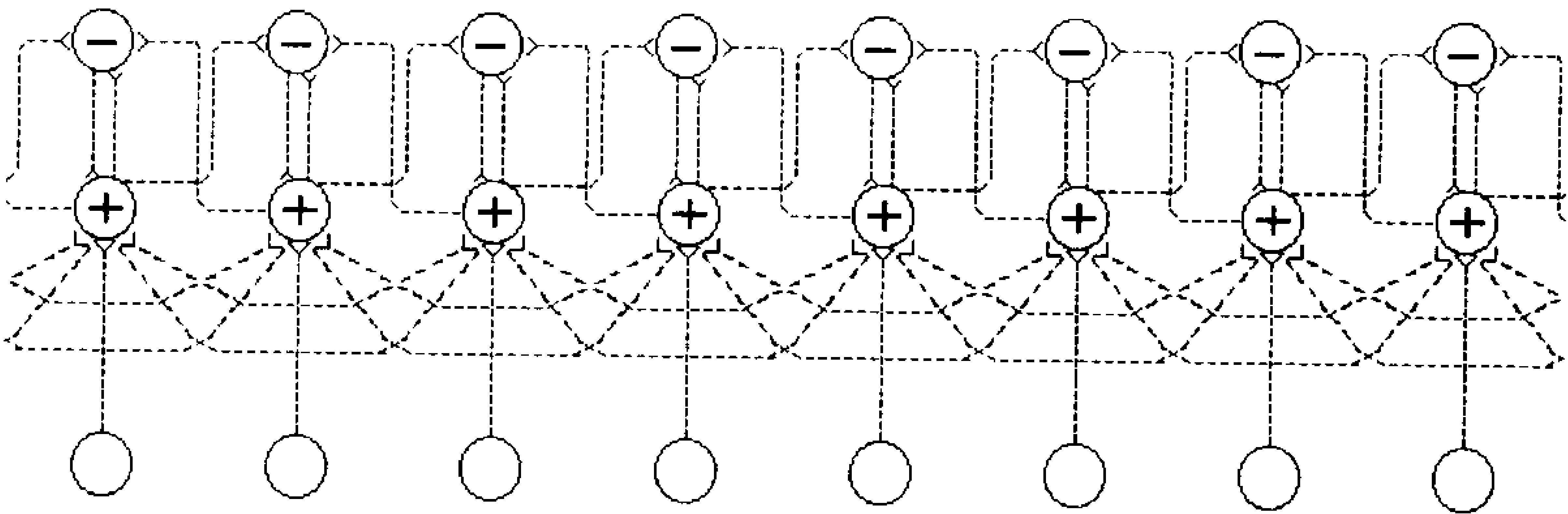


图 3.8.5　去同步化的线路联结 (说明详见正文)

(引自 Schillen, 1990)

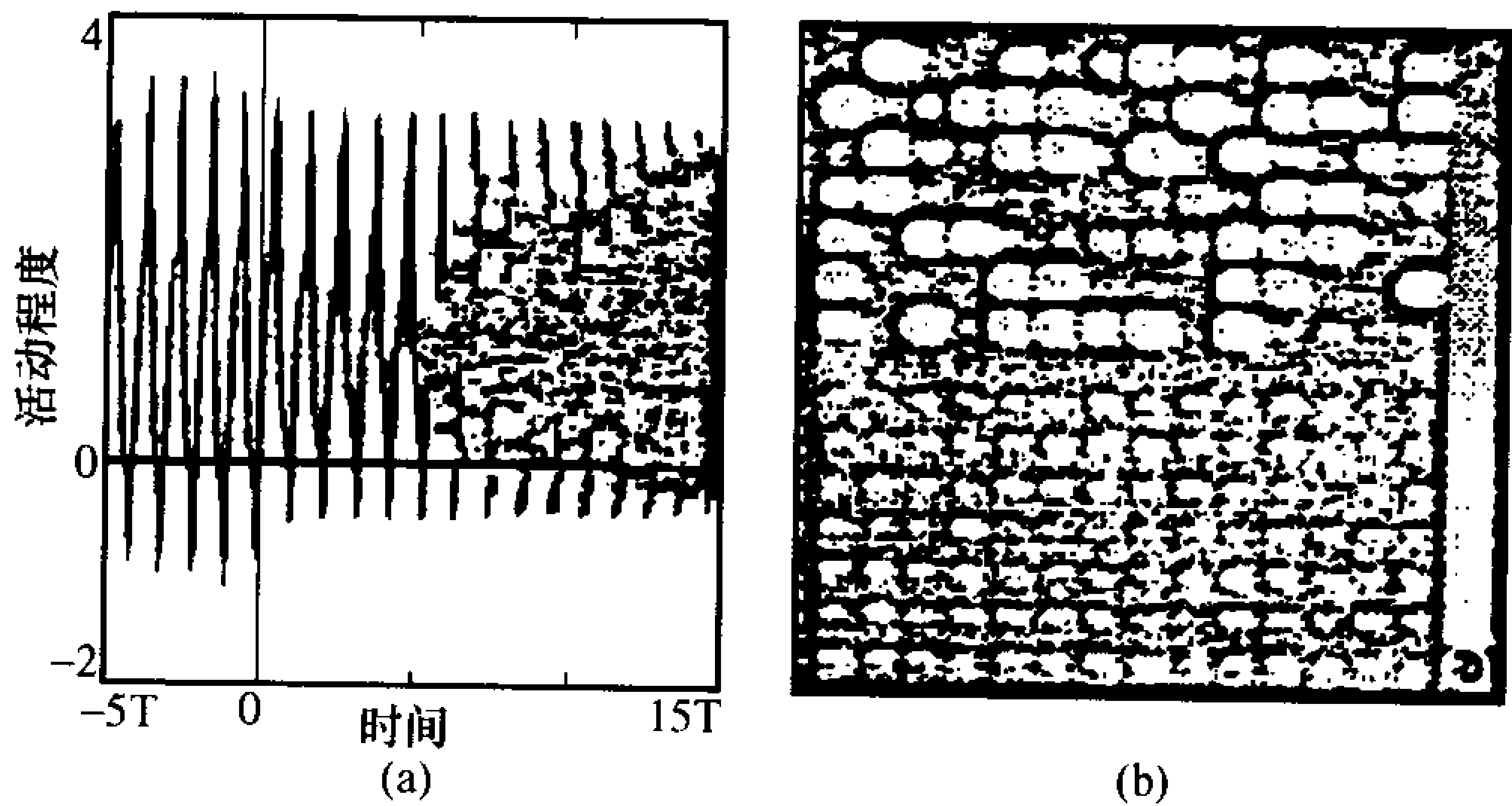


图 3.8.6　网络活动的去同步化 (说明详见正文)

(引自 Schillen, 1990)

把相邻单元的兴奋性单元和抑制性单元之间连接起来 (短程联结 - 同步化), 而把次相邻单元的兴奋性单元之间也连接起来 (长程联结 - 去同步化), 由这样构成的网络即可仿真上节中的生理学实验结果. 把 8 个振子按上述方式一维排列成一个链, 其中每个单元代表有不同最优朝向的细胞, 把带有阴影的方形表示在该处进行记录 (图 3.8.7(a) 上面的数字表示其最优朝向), 这些细胞的感受野彼此重叠. 图 3.8.7(b) 表示两种刺激条件: 左图只有一个光条刺激, 而右图则有两个光条刺激. 图 3.8.7(c) 表示在这两种刺激条件下每个细胞所受到的刺激强度. 图 3.8.7(d) 表示在这两种实验条件下, 由最优朝向分别为 112.5° 和 157.5° 的细胞处 (带阴影处), 分别记录到的活动. 图 3.8.7(e) 是对这两个细胞的活动所作的自相关

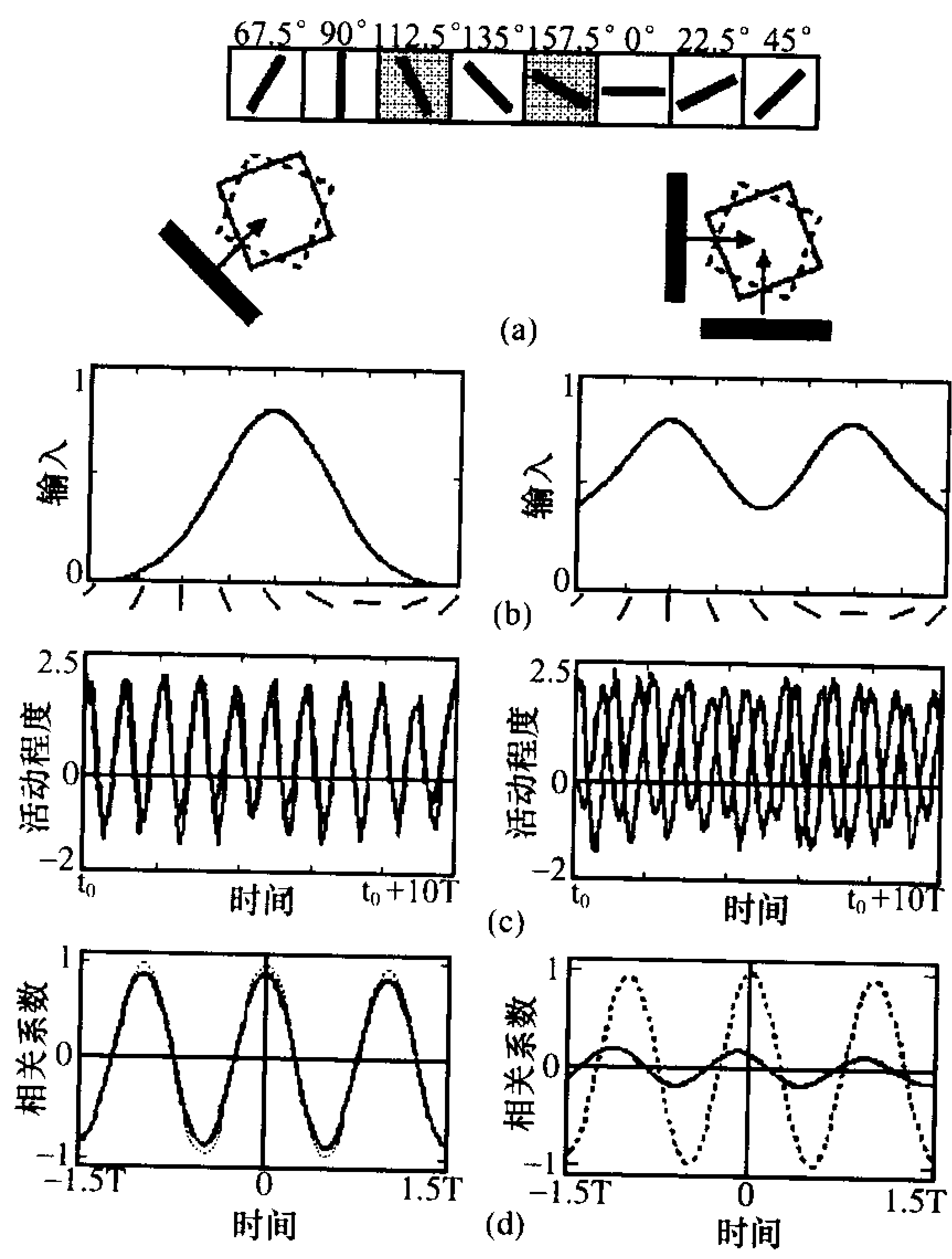


图 3.8.7　对于单元群体按刺激条件重组的仿真 (说明详见正文)

(引自 Schillen, 1990)

图 (虚线) 和互相关图 (实线). 由图 3.8.7 可见, 对于单光条刺激, 这两个细胞属于同一群; 而对于双光条刺激, 虽然这两个细胞的活动水平几乎维持不变, 但其互相关显著下降, 表明它们分别属于两个不同的群, 而检测两个不同的对象.

综上所述, 强烈的非线性是神经系统的突出特点, 这不仅表现为 S 形曲线式的静态非线性, 而且还表现为形形色色的动态非线性, 只研究固定点吸引子对生物神经网络来说是远远不够的. Freeman 指出, 对于脑电来说, 固定点吸引子意味着深度麻醉或死亡. Ronals 也指出对于诸如语言辨识或运动控制之类和时间有关的复杂任务来说, 必须考虑有奇异吸引子的动态系统. 由于问题的复杂性, 目前的工作尚属于起步阶段, 还有大量的问题有待我们去研究.

§3.9 动态神经元网络的混沌动力学性质

如果仅仅把神经元作为神经网络的一个元件来考虑, 那么就没有必要进一步深入到神经元的内部去考虑亚细胞水平的机制问题. 对于我们来说, 使我们感兴趣的仅仅是从信息处理的角度来说才有意义的神经细胞的主要外特性. 根据第一章中对神经细胞性质的介绍, 可以把这些主要特征归纳如下:

(1) 突触有两类: 兴奋性突触和抑制型突触;

(2) 存在突触延搁;

(3) 细胞膜有电阻 – 电容特性, 因而必然有时间总和特性;

(4) 有空间总和特性;

(5) 存在阈值, 当始段处膜电位达到阈值时就产生神经冲动;

(6) 发放后始段处膜电位迅速复极化甚或超极化;

(7) 发放后存在不应期: 绝对不应期和相对不应期, 亦即阈值并不是一个常数, 在发放后一段短时间内阈值上升到无穷大, 然后按指数衰减到静息水平;

(8) 动作电位服从全或无定律.

其中 (1)~(3) 属于神经元输入端 —— 突触的性质， (3)~(7) 属于神经元对输入信息进行整合之处 —— 始端的性质，而 (8) 属于神经元输出的性质.

如果我们暂不考虑不应期，另外如果神经元不是中间神经元或运动神经元，而是感应神经元，亦即它接收到外界刺激 $I(t)$ 而不是神经脉冲，那么这个外界刺激也可以在感觉神经元中产生感受器电流.

按照上述神经元输入端、整合部位和输出端的特性，可以分别建立它们的方程.

首先根据动作电位的全或无定律，可以近似地用 δ 函数来表示动作电位，因此第 j 个神经元的输出发生在一系列时刻 t_j^k $(k=1,2,\cdots,m)$ 的动作电位序列可以表示为

$$z_j(t)=\sum_k \delta(t-t_j^k).$$

考虑到膜的阻容性质，突触后膜电流 $x_{ij}(t)$ 应服从一阶微分方程，其中 $x_{ij}(t)$ 表示第 i 个神经元与第 j 个神经元所形成的突触处的突触膜电流. 假定此处的膜时间常数为 T_{ij}，则无疑可建立下列方程:

$$\frac{\mathrm{d}x_{ij}(t)}{\mathrm{d}t}=-\frac{1}{T_{ij}}x_{ij}(t)+A_{ij}z_j(t-\tau_{ij}), \tag{3.9.1}$$

其中 τ_{ij} 为此处的突触延搁， A_{ij} 为这个突触的作用系数，且符合决定突触的性质，兴奋性为正，抑制性为负.

对于感觉神经元来说，其输入如前所述已非神经脉冲序列而为外界刺激 $I_i(t)$. 如记感受器电流为 $x_{i0}(t)$，则类上可得

$$\frac{\mathrm{d}x_{i0}(t)}{\mathrm{d}t}=-\frac{1}{T_{i0}}x_{i0}(t)+A_{i0}I_i(t), \tag{3.9.2}$$

其中 T_{i0} 为膜时间常数， A_{i0} 为作用系数.

考虑始段处对局部电流的整合作用 (性质 (3)~(6))，若记始段

处分级电位为 $y_i(t)$, 则有

$$\frac{\mathrm{d}y_i(t)}{\mathrm{d}t} = -\frac{1}{T_i}y_i(t) + \sum_{j=0}^{n} w_{ij}x_{ij}(t) - \Theta_i z_i(t), \tag{3.9.3}$$

其中 T_i 为第 i 个神经元在始段处膜的时间常数, W_{ij} 为考虑到第 j 个突触在第 i 个神经元上的部位等因素在内所赋予的正的权重因子, Θ_i 是一个和阈值 θ_i 相等或稍大的数, 若是前者则表明有复极化, 否则则存在超极化.

最后, 当始段分级电位达到阈值 θ_i 时则引起发放脉冲, 因此其方程为

$$z_i(t) = \delta(y_i(t) - \theta_i). \tag{3.9.4}$$

方程 (3.9.1)~(3.9.4) 构成了能描述除了不应期之外神经元的上述 7 条性质的数学模型. 由于这个模型不把脉冲发放率这样的模拟量作为它的输出, 因此是一种脉冲发放神经元模型 (spiking neuron model), 而且由于在建立上述模型的过程中我们无须引进反映膜电流和脉冲发放率之间 S 形曲线这样的静态特性曲线, 因此把它称之为 **动态神经元模型**.

由于在上述模型中我们用了 δ 函数, 因此不能用传统的数值求解方法来计算或求解解析解. 然而, 若已知方程解在 t 时刻的值, 而要递归算得方程解在 $t+\Delta t$ 时的值, 则只要分别考虑在此区间内它接收到或产生脉冲与否两种情况, 就不难得出解此方程组的递归算法. 由于这个算法本身也有明确的生理意义, 因此也可把它看成是模型的离散形式. 此离散模型为

$$x_{ij}(t+\Delta t) = \begin{cases} \mathrm{e}^{-\frac{\Delta t}{T_i}} x_{ij}(t), & z_j(t-\tau_{ij}) = 0, \\ \mathrm{e}^{-\frac{\Delta t}{T_i}} x_{ij}(t) + A_{ij}, & z_j(t-\tau_{ij}) = 1. \end{cases} \tag{3.9.5}$$

若 T_{i0} 很小, 则近似有

$$x_{i0}(t+\Delta t) = A_{i0} I_i(t) \Delta t. \tag{3.9.6}$$

类似可得

$$y_i(t+\Delta t)=\begin{cases}\mathrm{e}^{-\frac{\Delta t}{T_{ij}}}\left(y_i(t)+\sum_{j=0}^{n}w_{ij}x_{ij}(t)\right), & z_i(t)=0,\\ \mathrm{e}^{-\frac{\Delta t}{T_{ij}}}\left(y_i(t)+\sum_{j=0}^{n}w_{ij}x_{ij}(t)\right)-\Theta_i, & z_i(t)=1,\end{cases}$$

(3.9.7)

$$z_i(t+\Delta t)=\begin{cases}1, & y_i(t)\geqslant\theta_i,\\ 0, & y_i(t)<\theta_i.\end{cases}\tag{3.9.8}$$

如果把 $z_i(t)=0$ 理解为第 i 个神经元在时刻 t 不发生脉冲，而 $z_i(t)=1$ 理解为第 i 个神经元在时刻 t 发生脉冲，则上述各式不仅给出了可直接在计算机上进行计算的算法，而且由于它已经取了离散形式，不存在将微分方程离散化可能产生原方程不存在周期解或混沌解的危险，这一点对于研究神经网络的混沌性质是至关重要的．否则即使得出混沌现象也将不能区分究竟这是网络所固有的性质，还是仅仅由于不恰当的离散化所带来的假象.

离散模型除了有上述优点之外，更便于加以修正以反映神经元的其它性质，如不应期之类．如要考虑不应期，只要假定阈值 θ_i 不再是常数，而是时间常数

$$\theta_i(t+\Delta t)=\begin{cases}M, & z_i(t)=1,\\ \mathrm{e}^{-\frac{\Delta t}{T_0}(\theta_i-\theta_{i0})+\theta_{i0}}, & z_i(t)=0,\end{cases}\tag{3.9.9}$$

其中 $M\gg\theta_{i0}$ 是一个很大的值， θ_{i0} 时第 i 个神经元的正常阈值(不落在不应期内)．这样，模型 (3.9.5)~(3.9.9) 就概括了上述所有 8 条主要性质.

利用上述模型，我们曾仿真了侧抑制网络的动态过程和稳态时的 Mach 带现象、水蛭中枢模式发生器的节律活动、耳蜗核中主

要的几种不同类型的神经元的刺激后时间直方图等等，仿真结果都和实验结果符合得很好，从而说明以上模型是在神经网络水平上进行模型和仿真研究的一个合适的出发点. 由于本书的主要目的在于讨论神经网络的非线性动力学性质，故对上述内容不拟多作叙述，而在下文中将只讨论由此类神经元模型所构造的某些神经网络的混沌性质.

关于脑电是否是混沌的这个问题自 20 世纪 80 年代以来一直是广泛地引起人们的兴趣而又存在争议的问题. 一方面从脑电所表现出来的不规则性，使人们易于猜测它是混沌的，由此人们利用各种混沌动力学指标如李氏指数、分数维等对其进行计算，计算所得的数值似乎支持脑电是混沌的设想. 然而由于脑电的极度非平稳性，而计算这些指标要求很大的数据点数这两者之间的矛盾，造成计算结果差异很大. 另外由于实际测量脑电都必须预先经过滤波，然而甚至是白噪声在经过适当的滤波之后才能表现出类似脑电这样的行为，因此关于脑电究竟是否是混沌的至今仍是悬案.

我们现在试图用动态神经元构成的网络对这一问题进行探索. 考虑到脑电是脑中大量神经元突触后膜电流活动的加权总和，脑中各神经元之间既存在着兴奋性突触联结，又存在着抑制性突触联结. 由于突触联结数目巨大，我们可假定这种联结的权重服从某种统计分布. 由于 $x_{ij}(t)$ 具有突触后膜电流的意义，在没有更详细的知识之前，我们试探性地初步用量 $\sum\limits_{i,j} x_{ij}(t)$ 作为近似来代表一个在某种意义上类似于脑电的量，用以模拟产生脑电的整个脑神经网络. 我们假定其权重联结服从在某个有正有负的范围内的均匀分布. 当然这些假设从定量上讲和真实脑的情况相差甚远，因此不能期望由此所产生的 $\sum\limits_{i,j} x_{ij}(t)$ 在细节上会和实验记录到的脑电一模一样，只是想考察即使在这样非常粗略的近似之下，$\sum\limits_{i,j} x_{ij}(t)$ 是否会表现出某种混沌性质. 如果是的话，那么这将说明把具有生物神经元主要外特性的元件按既有兴奋性联结又有抑制性联结且以权重服从某种统计分布的方式彼此联结起来，就会产生某种混

沌动力学行为，那么在更复杂如真实脑这样的情况之下自然更有可能产生具有混沌性质的脑电了.

图 3.9.1(a) 是在零初始条件之下计算所得的结果. 它确实表现出和脑电有些相似的无规则行为. 图 3.9.1(b) 是对初始条件作微小改变之后计算所得的结果. 虽然用肉眼很难区别这两者之间的差别，然而如果从图 3.9.1(a) 中的曲线减去图 3.9.1(b) 中的曲线，则其结果如图 3.9.1(c) 所示. 由图可见两者之差在开始 100ms 内几乎为零，然而在 100ms 以后，其差别越来越大，且无规律可循. 如果按照 Rapp 关于混沌的三大特点来加以考察：(1) 系统必须是确定性的；(2) 表现出不规则的行为；(3) 对初始条件极端敏感. 那么这样的结构所得的神经网络就是一个混沌神经网络. 要指出的一点是在上述模型中仅在各神经元的外界输入端上加上均匀的恒定输入偏置 I，联结权重虽然是随机给定的，但一旦给定以后，它就不再改变. 因此在上述结构中并不存在任何随机干扰，然而它

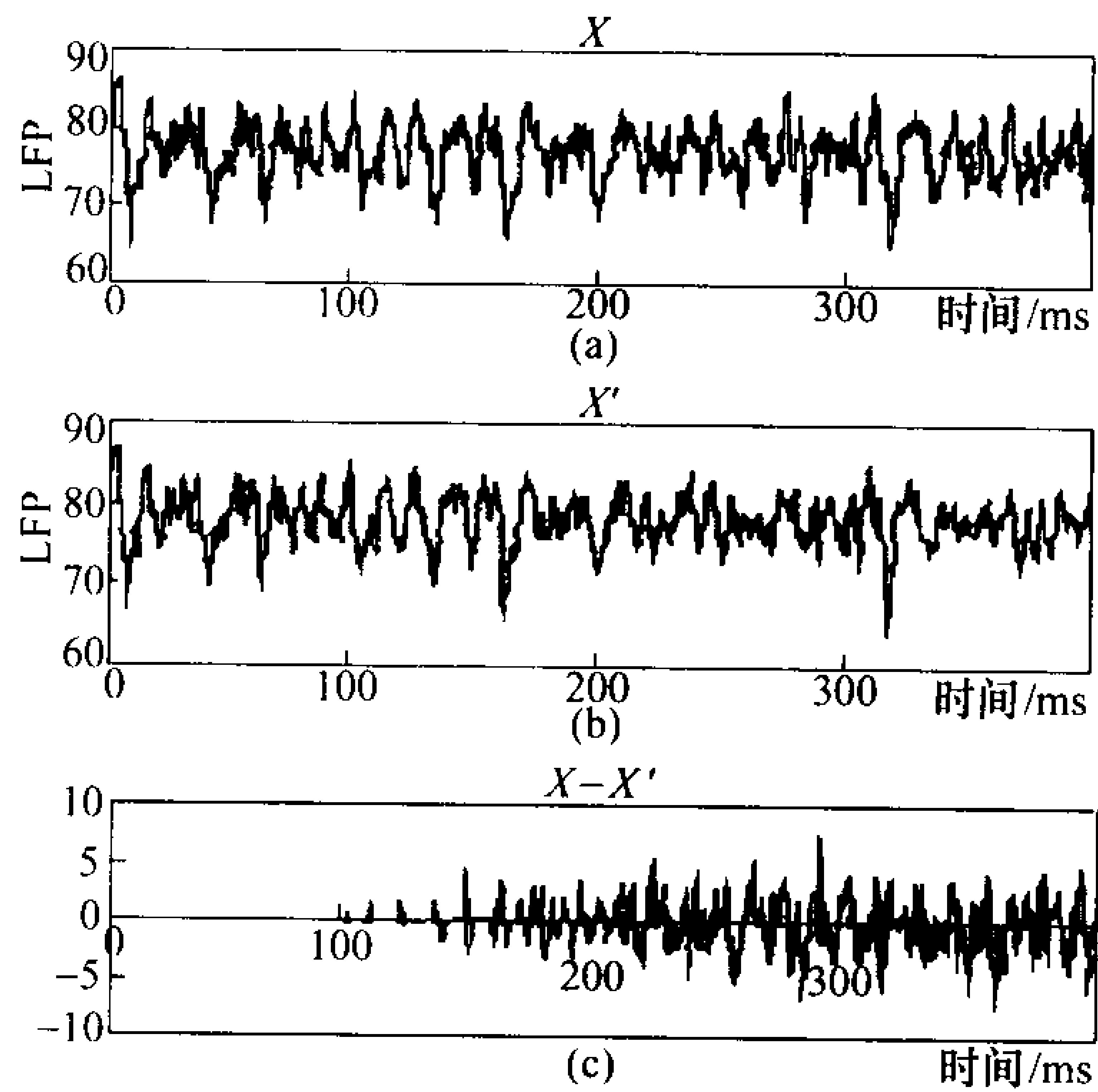


图 3.9.1　(a) 在零初始条件之下，权重服从 $[-0.7, 0.3]$ 之间均匀分布，各神经元加上恒定均匀偏置的神经网络的类脑电输出 $\sum_{i,j} x_{ij}(t)$；(b) 当初始条件略加变化时同一网络的同一输出；(c) 图 (a) 和图 (b) 之差

却依然表现出对初值极端敏感的不规则行为. 这也许是从一个侧面对脑电混沌说的支持.

与此同时, 我们还试图从方程 (3.9.3) 本身来分析动态神经元网络的混沌学性质. 在文 [40] 中, 我们考虑了经过简化的一般模型

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x), & t \neq \tau_k, \quad k = 0, 1, 2, \cdots, \\ \Delta x = I_k(x(t)), & t = \tau_k, \quad k = 0, 1, 2, \cdots, \\ x(0+) = x^*. \end{cases} \tag{3.9.10}$$

上述形式是对带有脉冲的微分方程的一种标准写法, 其中 $f(x)$ 和 $I_k(x)$ $(k = 0, 1, 2, \cdots)$ 均是定义在 $D \subset \mathrm{R}^n$ 上的函数, 而且 $0 = \tau_0 < \tau_1 < \cdots < \tau_k < \cdots$, $\tau_{k+1} - \tau_k = T$, $\lim\limits_{k \to +\infty} \tau_k = +\infty$, $\Delta x = x(t+) - x(t)$.

首先假设如下常微分方程 (组) 的初值问题

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = f(x), \\ x(t_0+) = x_0 \end{cases} \tag{3.9.11}$$

对任意的 $x_0 \in D$ 及 $t_0 \in [0, +\infty)$, 存在唯一的解 $\phi(\cdot; t_0, x_0) : (t_0, \beta) \to \mathrm{R}^n$, $\beta - t_0 \geqslant T$, 并且这一解是对初值连续依赖的, 于是有如下结论:

假设脉冲微分方程 (3.9.10) 的脉冲具有形式

$$I_k(x) = y_k - x, \quad t = \tau_k, \quad k = 0, 1, 2, \cdots, \tag{3.9.12}$$

其中序列 $\{y_k\}$ 满足 $y_{k+1} = g(y_k)$, $k = 0, 1, 2, \cdots$, 那么当映射 $g : Y \to Y \subset D$ 满足 Devaney 意义下混沌的条件时, 则系统 (3.9.10) 亦是在 Devaney 意义下混沌的.

事实上对上述结果的证明，可以考虑构造 Poincaré 栅栏 (截面) 上的离散映射. 令 $B_k = \{(t,x) \mid x \in D, t = \tau_k\}$ $(k = 0,1,2,\cdots)$ 为空间 t–x 中的 Poincaré 栅栏，于是根据脉冲微分方程 (3.9.10)，有如下 Poincaré 映射:

$$P = \psi \circ g \circ \psi^{-1} : V \to V,$$

其中 $\psi(x) = \phi(\tau_k + T; \tau_k, x) = \phi(T; 0, x)$, $V = \psi(Y)$. 由于我们对解 ϕ 的性质作了良好的假设，所以映射 ψ^{-1} 是存在的. 我们还知道如果一个连续系统导出的 Poincaré 映射是混沌的，则该系统本身亦是混沌的，所以下面要证明的是映射 P 在 Devaney 意义下是混沌的. 所谓 Devaney 意义下的混沌需要满足以下 3 条性质:

(1) 对初值的极端敏感性;

(2) 拓扑传递性;

(3) 周期轨道的稠密性.

因此，首先要证明映射 P 对初值是极端敏感的，也就是对任意适当的 $\varepsilon > 0$, 及 $x_0, \overline{x}_0 \in V$, 存在一个 $N > 0$, 使得

$$|P^N(x_0) - P^N(\overline{x}_0)| > \varepsilon.$$

因为 $P = \psi \circ g^N \circ \psi^{-1}$, 于是有

$$|P^N(x_0) - P^N(\overline{x}_0)| = |\psi[g^N(y_0)] - \psi[g^N(\overline{y}_0)]| > \varepsilon,$$

其中 $y_0 = \psi^{-1}(x_0)$, $\overline{y}_0 = \psi(\overline{x}_0)$. 以下用反证法.

如果以上叙述不成立，即存在一个 $\varepsilon_0 > 0$, 对所有的 $n \geqslant 0$, 使得

$$|\psi[g^n(y_0)] - \psi[g^n(\overline{y}_0)]| \leqslant \varepsilon_0.$$

然而，由于函数 $\psi(x)$ 关于 x 是连续的，则存在 $\delta_0(\varepsilon_0) > 0$, 使得

$$|g^n(y_0) - g^n(\overline{y}_0)| < \delta_0, \quad 对任意的\ n \geqslant 0.$$

但是这与混沌映射 g 的性质是矛盾的. 所以映射 P 是对初值极端敏感的.

其次，我们要证 P 是拓扑传递的，即对任意的开集 $E, F \subset V$，存在 $N > 0$，使得

$$P^N(E) \cap F \neq \emptyset.$$

我们知道对 $\psi^{-1}(E), \psi^{-1}(F) \subset Y$，存在 $N > 0$，使得

$$g^N[\psi^{-1}(E)] \cap \psi^{-1}(F) \neq \emptyset.$$

上式两边作用 ψ 得到

$$\psi\{g^N[\psi^{-1}(E)]\} \cap F \neq \emptyset,$$

注意到

$$\psi g^N \psi^{-1} = \psi g \psi^{-1} \psi g \psi^{-1} \cdots \psi g \psi^{-1} = P^N,$$

所以我们证明了映射 P 具有 Devenay 意义下的第二条性质.

最后，因为方程 (3.9.11) 的解 ϕ 是唯一的，所以映射 ψ 是一对一的. 又根据 $\overline{\mathrm{per}(g)} = Y$ 与 $\psi(Y) = V$，那么 $\overline{\mathrm{per}(P)} = V$ 是显然的.

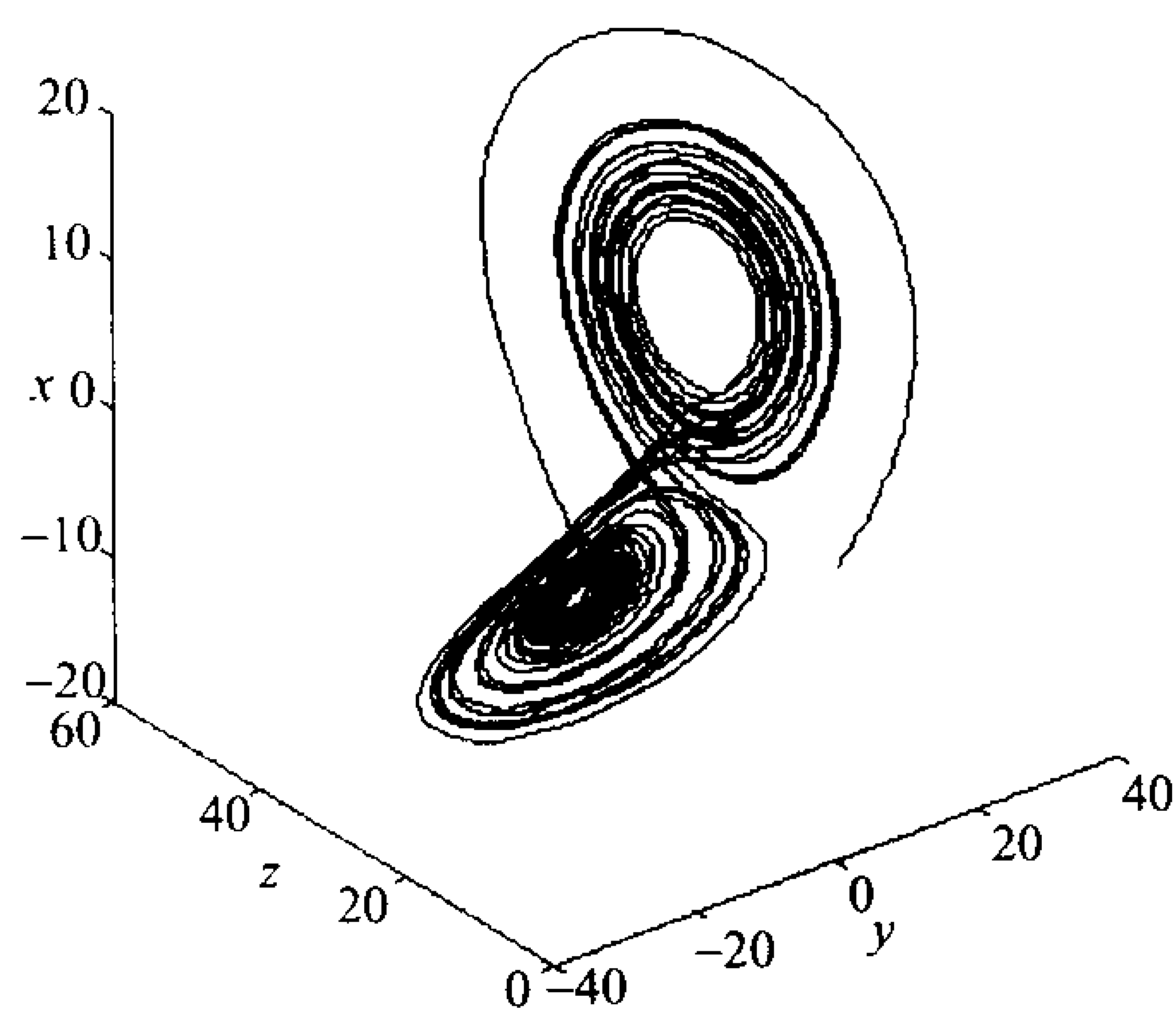
综上所述，系统 (3.9.10) 是 Devaney 意义下混沌的.

下面给出一个具体的实例来说明上述结论. 考虑如下的三维 Lorenz 系统:

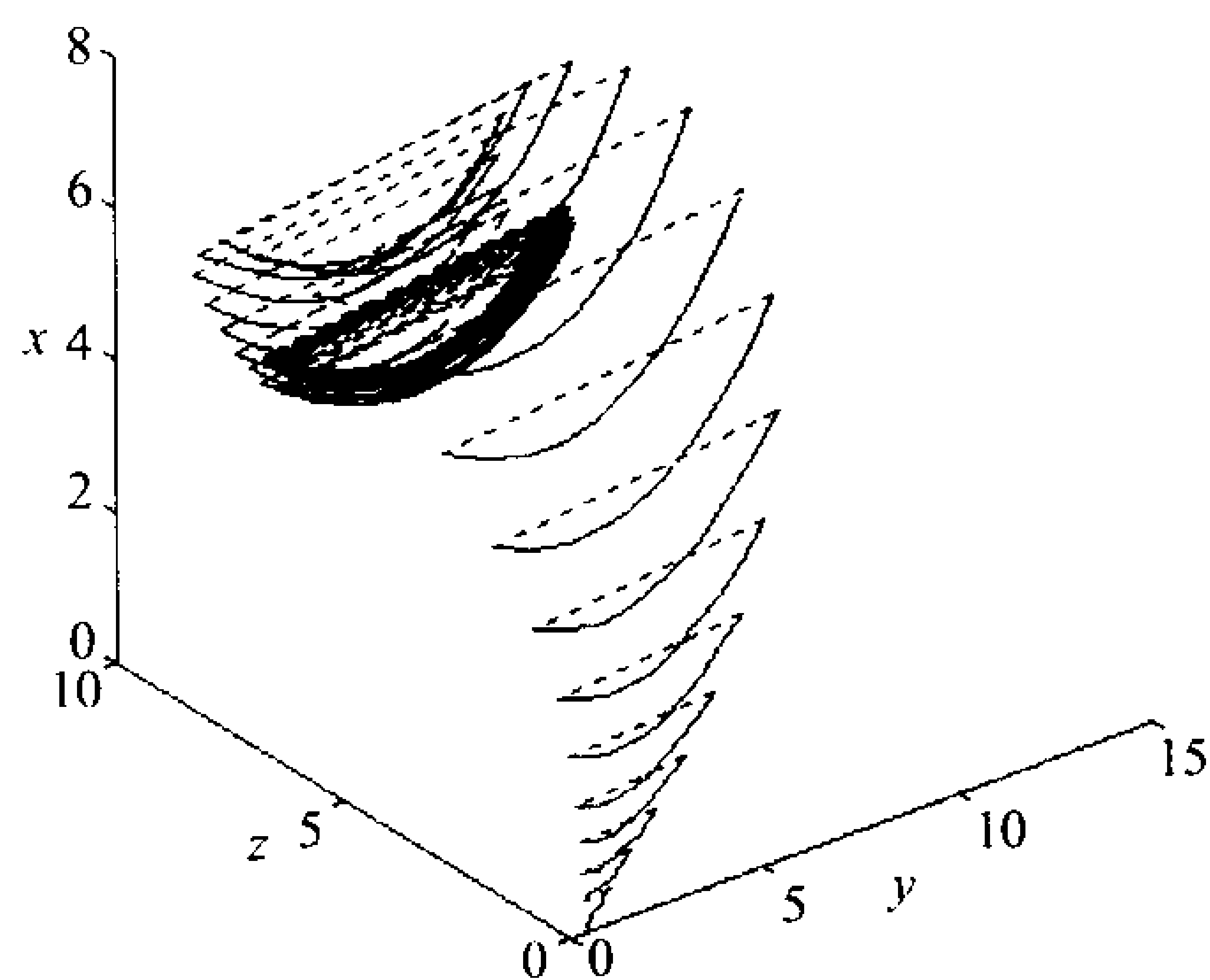
$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = -\sigma x + \sigma y, \\ \dfrac{\mathrm{d}y}{\mathrm{d}t} = -xz - y + \gamma x, \\ \dfrac{\mathrm{d}z}{\mathrm{d}t} = xy - bz. \end{cases} \tag{3.9.13}$$

已知当参数 $\sigma = 10, \gamma = 28, b = \dfrac{3}{8}$ 时，系统 (3.9.13) 是混沌的. 然而我们可以通过引进脉冲后，(3.9.13) 既可以变化为稳定的，也可以逐渐演化出新的混沌状态. 如当引进脉冲后的微分方程有如下形式:

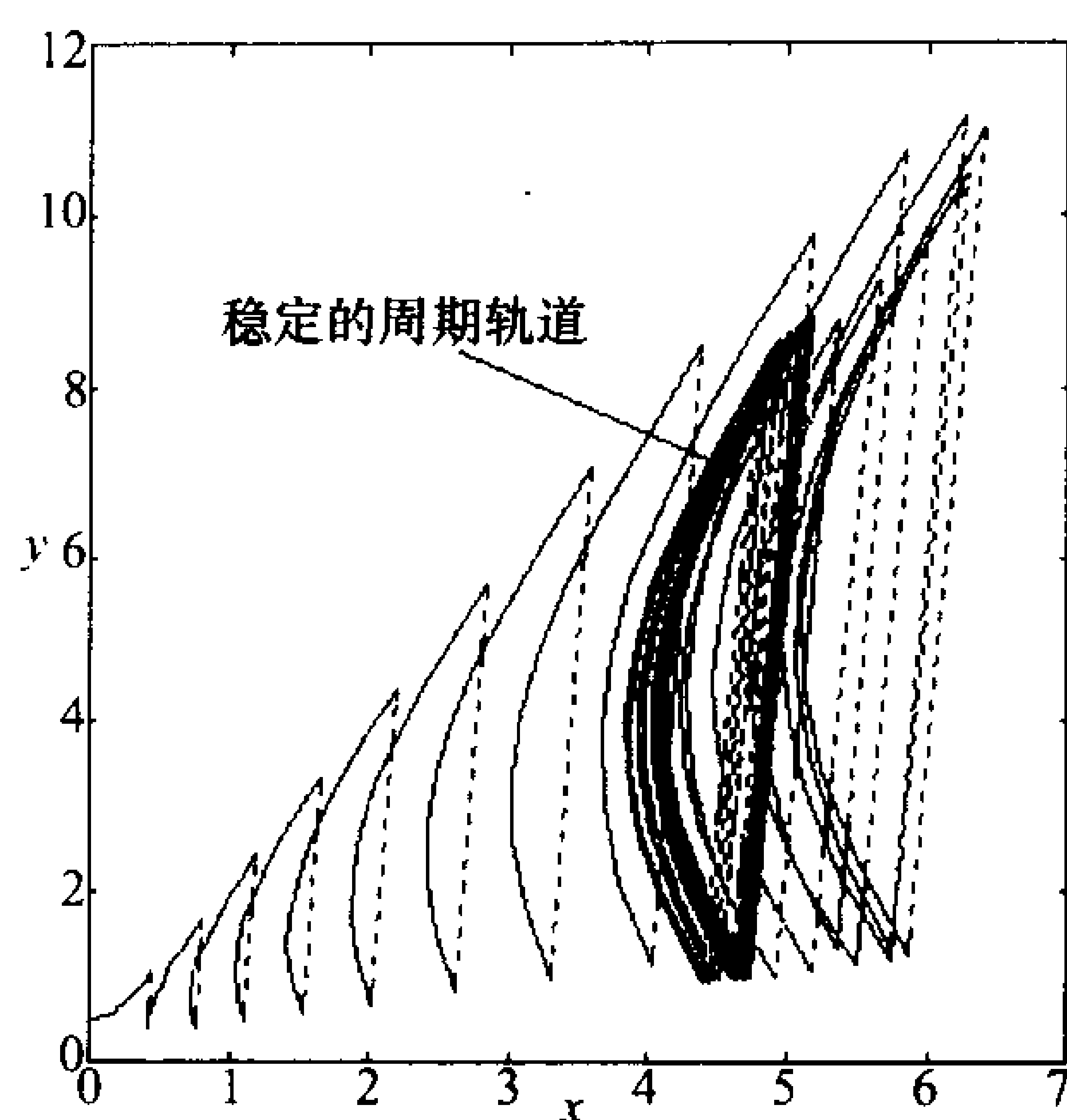
$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t} = -\sigma x + \sigma y, \\[2ex] \dfrac{\mathrm{d}y}{\mathrm{d}t} = -xz - y + \gamma x, \quad t \neq 0.1k, \quad k = 0, 1, 2, \cdots, \\[2ex] \dfrac{\mathrm{d}z}{\mathrm{d}t} = xy - bz, \end{cases} \tag{3.9.14}$$



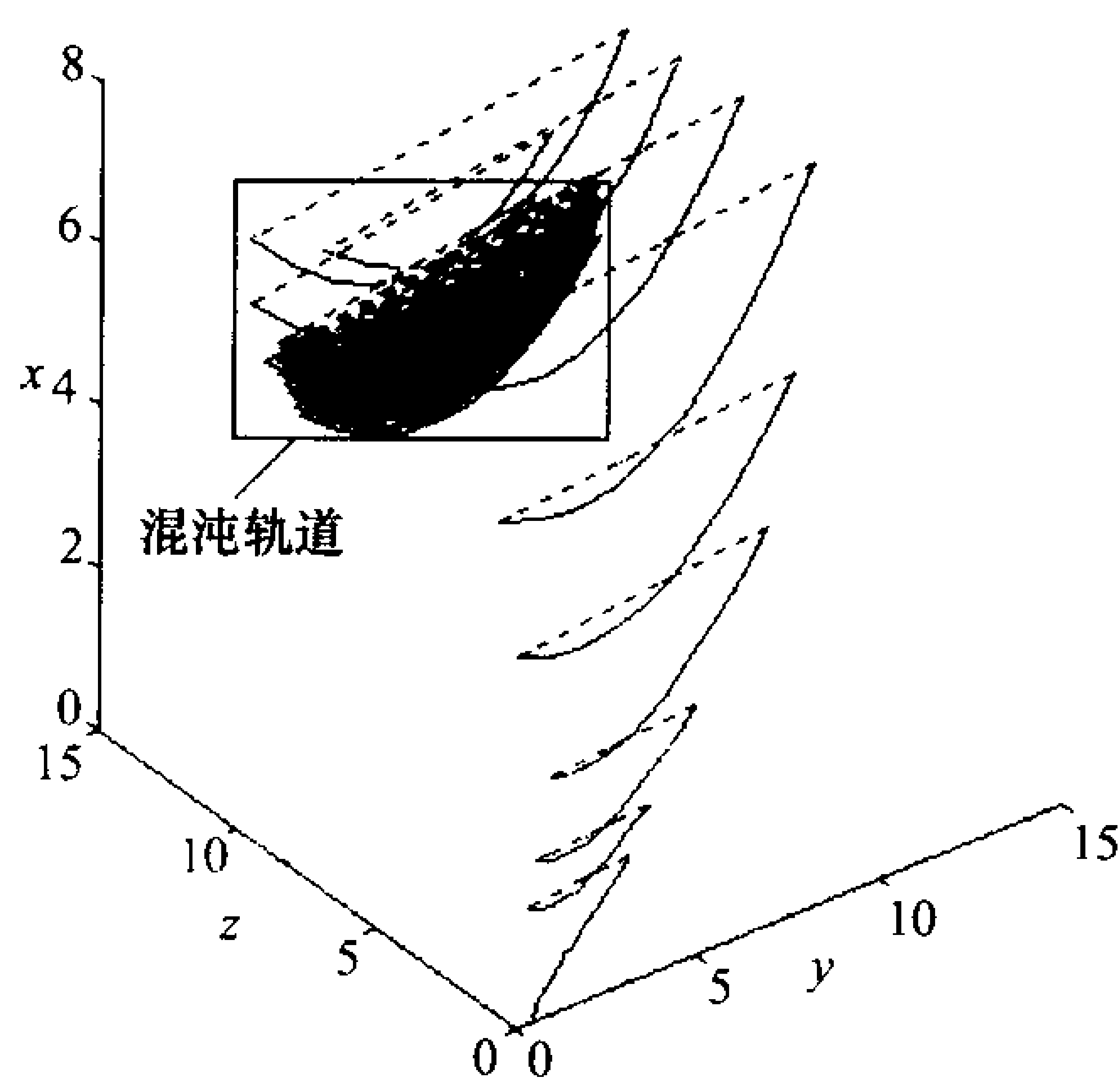
(a)



(b)



(c)



(d)

图 3.9.2　方程 (3.9.13) 和方程 (3.9.14) 的状态演化

(a) 混沌：方程 (3.9.13) 在相空间 x–y–z 中的图像；(b) 由混沌到有序：当 $\lambda = 1.000$ 时，方程 (3.9.14) 在相空间 x–y–z 中的轨线；(c) 由混沌到有序：当 $\lambda = 1.000$ 时，方程 (3.9.14) 在相平面 x–y 中的轨线；(d) 由混沌到混沌：$\lambda = 3.900$ 时，方程 (3.9.14) 在相空间 x–y–z 中的轨线

$$
\begin{cases}
\Delta x = I_k^1(x) = 0, & t = 0.1k, \quad k = 0, 1, 2, \cdots, \\
\Delta y = I_k^2(y) = y_k - y, & t = 0.1k, \quad k = 0, 1, 2, \cdots, \\
\Delta z = I_k^3(z) = 0, & t = 0.1k, \quad k = 0, 1, 2, \cdots, \\
y_{k+1} = g(y_k) = \lambda y_k(1 - y_k), & y_0 = 0.5.
\end{cases}
$$

那么图 3.9.2 给出了当参数 λ 取不同数值时，系统 (3.9.14) 状态的演化过程.

当然上述的结论也只是对深入讨论动态神经元网络混沌行为做了一些理论上的准备，更深层次的探讨工作正在进行之中.

§3.10　具有延迟的神经网络动力学系统模型和问题

3.10.1　嗅觉皮层的具延迟神经动力学模型

为了研究皮层震荡动力学行为和联想记忆之间可能存在的联系，普通认为嗅觉系统在许多方面都是一个极好的模型系统. 现通过利用简单的网络单元和现实结构，Hans Liljenströn (1991, 1992) 建立了关于嗅觉皮层的一种抽象模型，它具有中等的复杂程度. 网络单元相应于带有连续输入 - 输出关系的神经元群体，描述了脉冲密度特征. 所有的单元根据关于嗅觉皮层结构的已知事实相连接，即依据 Halerly (1985) 所决定的基本电路. 这意味着一个带有两层抑制单元和一层激励单元的三层结构 (见图 3.10.1). 顶层包括抑制的 “前馈内神经细胞”，它从外部源 (“嗅觉泡”) 和激励 “锥细胞” 中间层来接受输入，它们又是局部地投射到激励单元. 底层包括抑制的 “后馈内神经细胞”，又从激励单元接受输入并且投射回激励单元去. 这两个抑制单元的集合由两个不同的时间常数和与激励单元有些不同的联系而表征. 除了从抑制单元得到的反馈，激

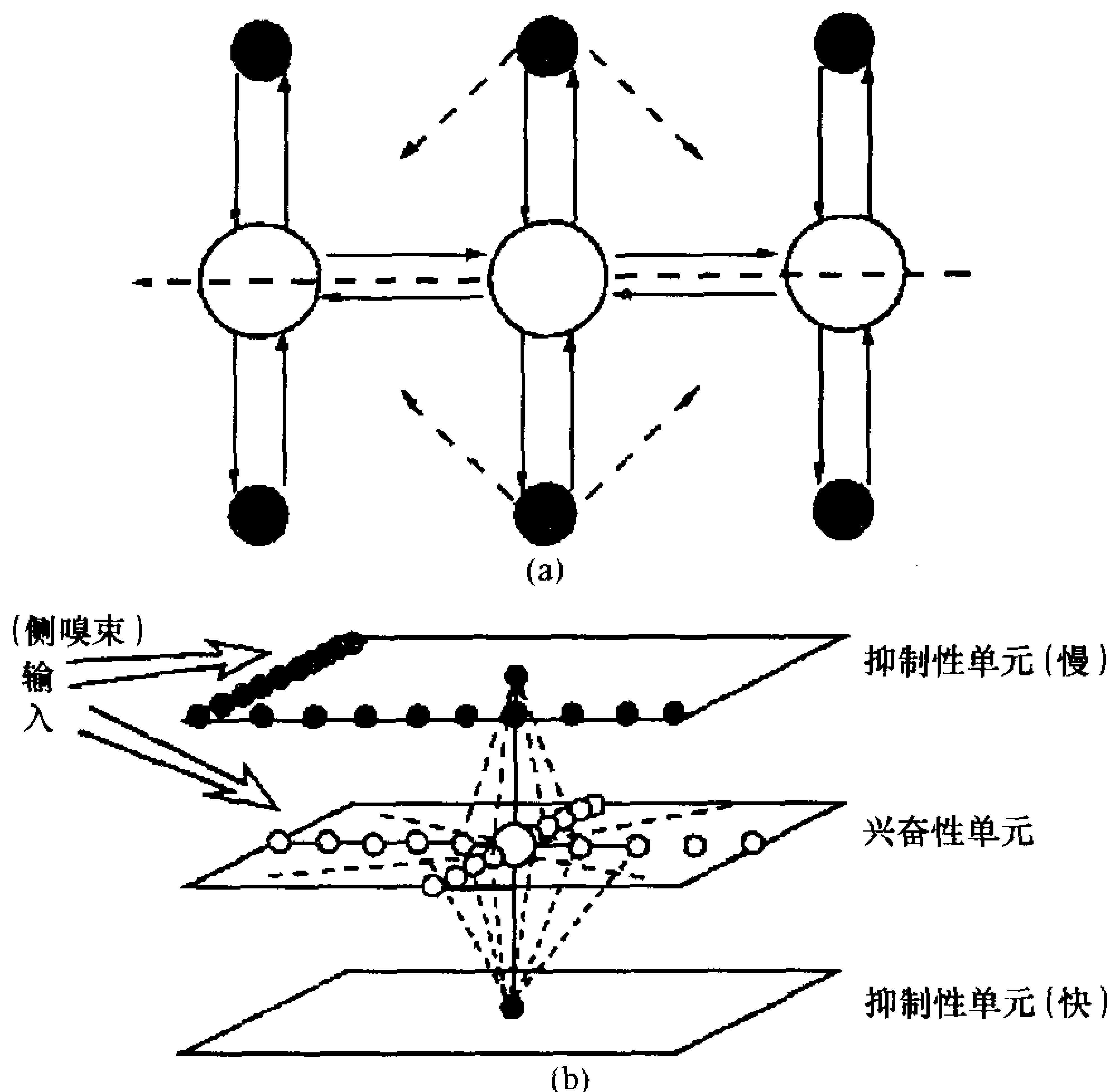


图 3.10.1　三层模型的嗅觉皮层的神经系统

励单元彼此之间以及从"嗅觉泡"都接受大量的输入，所有的连接同信号传播的时间时滞一起形成模型，相应于实际皮层中的几何和结构特征.

对每一个单元 I, 其行为状态变量为 u_i, N 个神经网络随时间演化的动力学方程，有外部输入 $I(t)$, 特征时间常数 τ, 单元 i 与 j 之间的连接权重 w_{ij}, 时间延滞 τ_i, 输入输出函数是一个连续的 Sigmoid 函数，由 Freeman (1979) 实验决定.

该网络模型可以写成以下的具延迟的微分方程组:

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \sum_{j=1}^{N} w_{ij} g_j [u_j(t - \delta_{ij})] - \frac{u_i}{\tau_i} + I_i(t), \tag{3.10.1}$$

$$g_j(u_j) = cQ_j \left\{ 1 - \exp\left[-\frac{\mathrm{e}^{u_j} - I_j}{Q_j} \right] \right\}, \tag{3.10.2}$$

其中有常数为每一个激励单元的增益常数 Q_i, 依赖于那个特定单

元的前面的行为，这可以作为由于细胞中钙的积累而产生的实验上观测的神经适应的一种近似，还有一个适应参数 D. 而 Q_c 是决定最大增益的常数，$\langle u_j(t)\rangle_{\tau_0}$ 是沿着最后 τ_0 秒的单元行为的时间平均，神经调节的影响通过改变 Q_c 或 D 而得以模拟

$$Q_j = Q_c \exp[-\langle u_j(t)\rangle_{\tau_0}^2/D]. \tag{3.10.3}$$

上述延迟系统中状态变量 u 与非线性输出函数 g 以及增益常数 Q 及参数 D 之间的关系如图 3.10.2 所示.

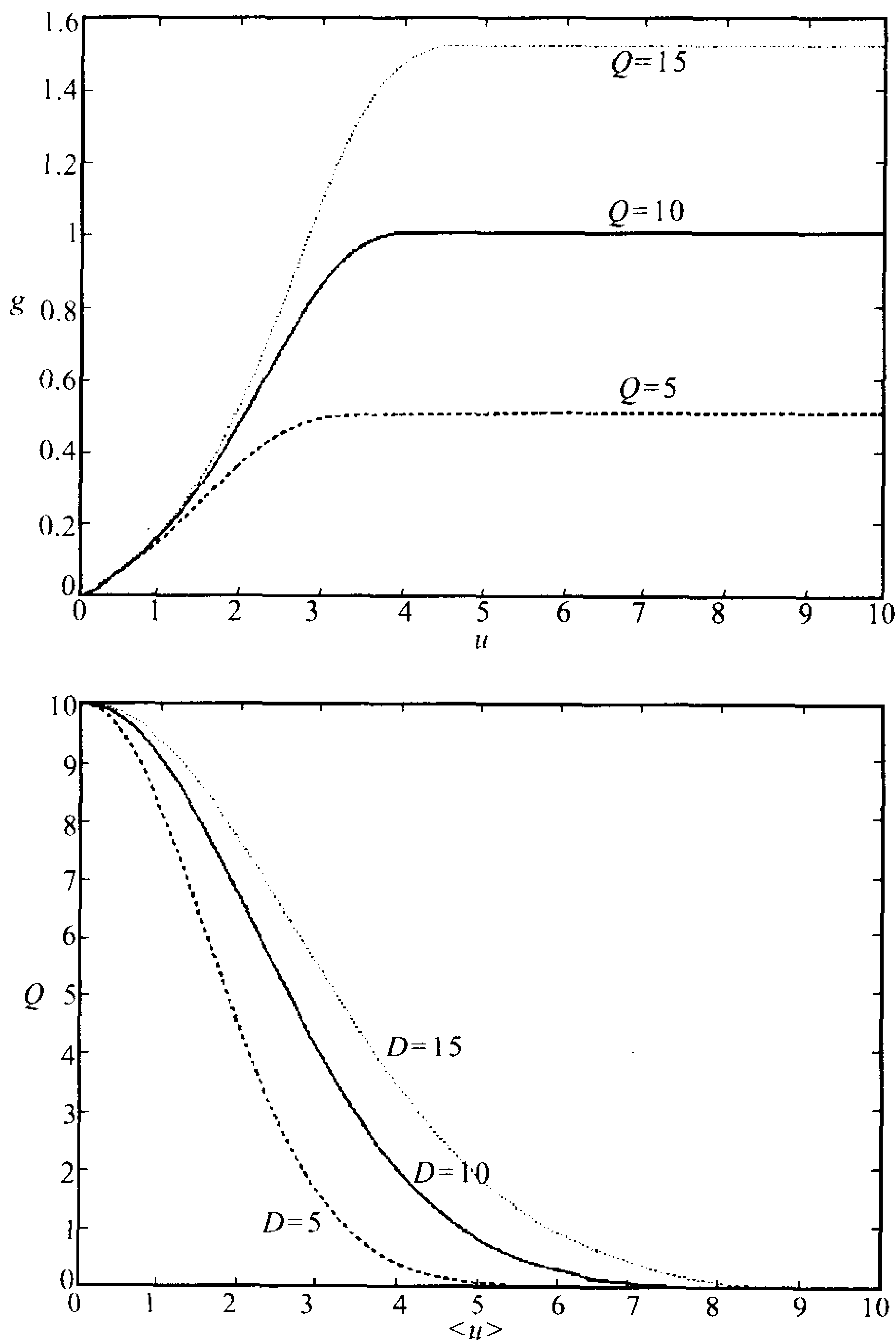


图 3.10.2　g 与 u, Q; Q 与 $\langle u\rangle$, D 相依关系图

外部输入 $I_i(t)$ 随现实时间延滞赋予每一个激励单元以及前馈式抑制单元, 模拟了从嗅觉泡通过 LOT 到皮层的传入的输出 (LOT 为横向嗅觉管). 大多数情况下, 输入是随机的, 或者是常数或者随时间变化. 当模拟振动的 LOTS 脉冲时, 所有的输入"结构 (纤维)" 将以两个模拟毫秒的高振幅运动.

为了能够进行学习和联想记忆, 连接权 w_{ij} 根据某些学习规则递增变化, 选择几种学习规则来考察其特定的动力学行为. 特别地, 考虑一种传导延迟 δ_{ij} 处于一个网络单元的输出行为和它在接收单元的效果之间的情形. 连接权重强度的变化也依赖于那个连接权重的绝对值, 以至于这个值不能超过某些最大权重强度 $w_{\max}$. 对于学习效率 η, 在时刻 t 单元 j 和单元 i 之间连接权重的改变由下式给出

$$\Delta w_{ij} = \eta g_i(u_i) g_j[u_j(t-\delta_{ij})](w_{\max} - w_{ij}). \tag{3.10.4}$$

当模拟突触传导的神经调节效果时, 改变权常数 w_c, 它会随着印有激励连接权重而加倍, 可用 100 个激励单元与 200 个抑制单元 (共 300 个) 对皮层神经调节联想记忆动力学进行计算机模拟.

3.10.2 振荡和混沌的嗅觉皮层动力学行为的模拟

延迟模型 (3.10.1)~(3.10.3) 的模拟计算能够重现实际嗅觉皮层的主要特征动力学行为. 特别地, 由模型可以同时展示 5Hz 和 40Hz 的振荡, 这是对实际皮层的 θ 节奏和 γ 节奏的模拟. 又对强弱震荡及对 LOT 产生的效果进行了模拟, 对前者产生了双相的反应, 后者是持久的振荡反应. 这些效果作为行为波像海波一样沿网络扩展, 强脉冲导致了一个沿表面移动的快波, 而脉冲产生了几个具有衰减振幅的连续波. 所有这些现象都关键地依赖网络结构, 尤其是前馈和反馈抑制环以及瞬时延迟下的长程激励连接.

在某些条件下, 系统表现出类似混沌的行为 (见图 3.10.3). 这些动力学行为类似于 EEG 中所看到的. 此外还模拟了带有神经适应和神经调节作用的振荡动力系统, 通过在模型中引进神经适应, 即使在不涉及抑制电路的情况下也可能获得振荡行为. 在这种

情况下振荡频率依赖于适应参数 D 和激励单元之间的连接强度. 特别地，当前依赖于模型前馈抑制单元的时间常数的低频振荡，现在可以由这种适应作用而得以解释. 在没有外部输入时，这样的振荡可以由系统的内部噪声得到 (见图 3.10.4).

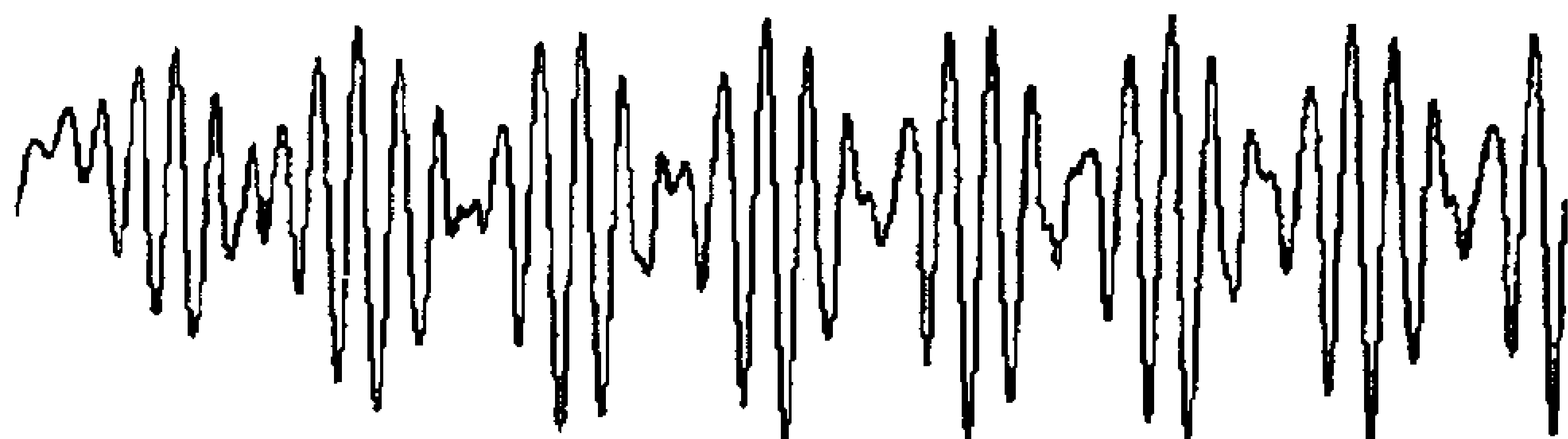


图 3.10.3　延迟模型 (3.10.1)~(3.10.3) 中的类似混沌行为的图

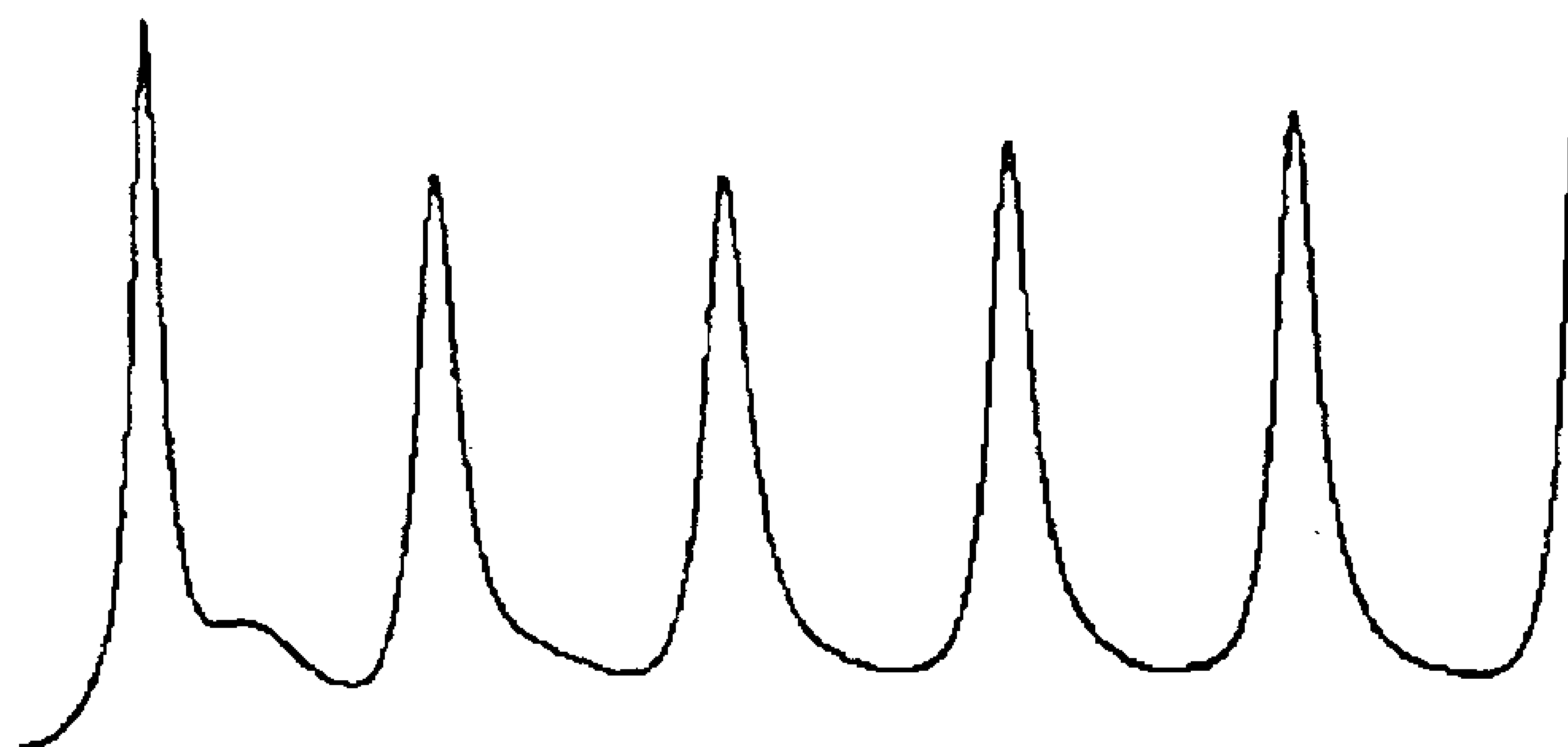


图 3.10.4　没有外部输入时的系统振荡图

对于延迟模型模拟学习和联想记忆中的神经调节效用结果表明，那样的效果可以导致快速学习和记忆，或者是一种新的输入模式和预先储存模式的更精确的联系，反映了时间和精确性之间有一种交替. 可以想象这种交替在当前情况下由更高级的皮层结构所决定. 记忆过程中的神经调节水平，相应于唤起状态，也能决定问题所必需的精确程度.

利用前面提到的 Hellian 学习规则的更改模式，还训练了带有振荡或非振荡的输入模式的网络. 振荡输入将相应于从嗅觉泡通过 LOT 传送到嗅觉皮层的味觉信息类型. 这种模式作为极限环吸引子状态而储存起来，即在记忆恢复时，回忆模式将以 r 频率振

荡. 有神经调节效用的学习比没有神经调节效用的学习要快, 但是主要的效果是回忆时间和精确度, 带有记忆干扰抑制的一种更有效的学习使得利用任何储存模式去匹配一种降级的输入模式具有更高的精确性成为可能 (见图 3.10.5).

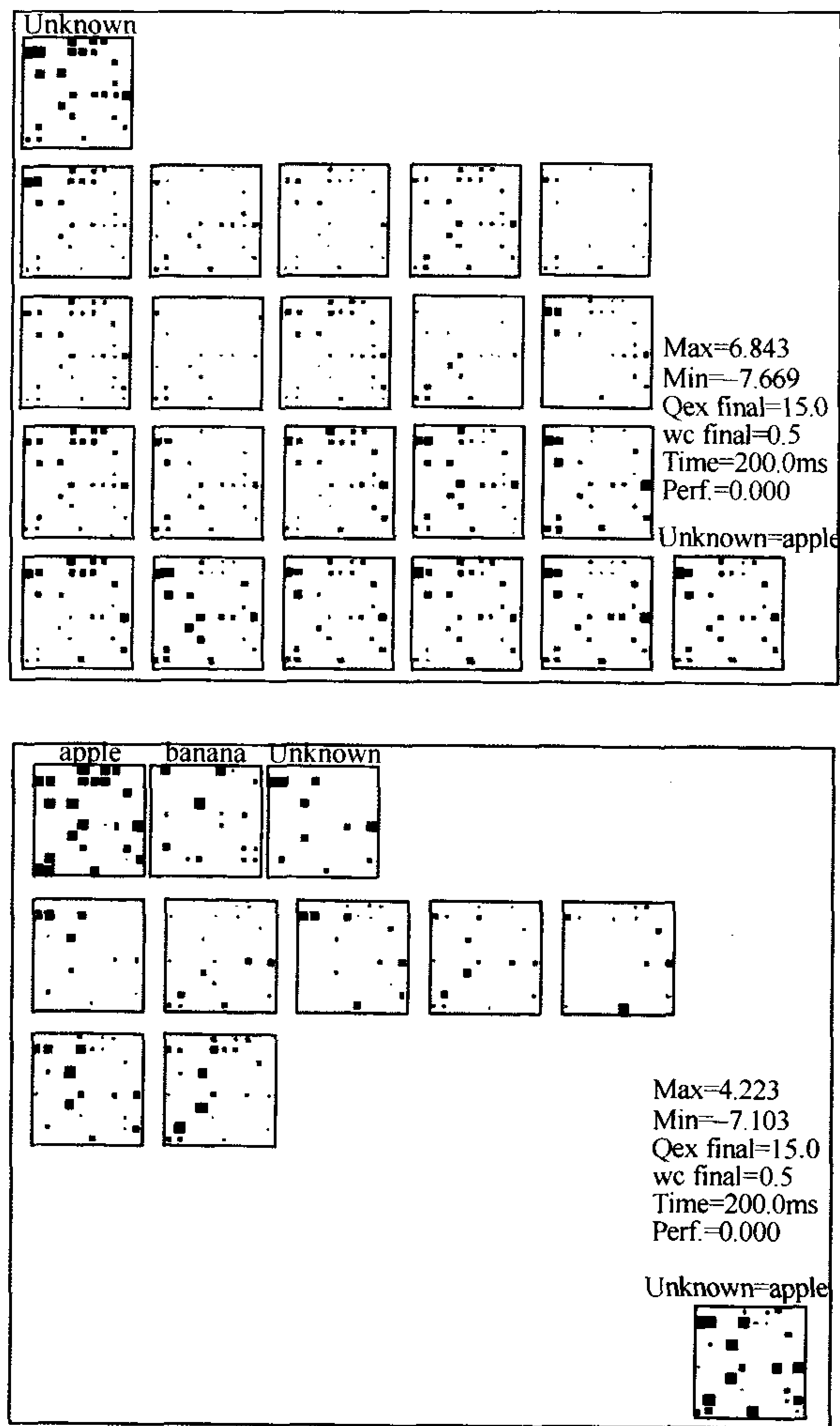


图 3.10.5 延迟模型学习和联想记忆的数值模拟图

在嗅觉皮层以及前述的模型中模拟所观察到的高度非周期和不规则行为可以在某种意义下归入“混沌”一类，但通常需要更长的时间序列来决定它的真正本性. 可以想象，当系统在两个不同的极限环或吸引子记忆状态之间移动时，混沌状态又是暂时的过渡，即周期解与混沌解的混合及阵发性，从而反映了记忆和思维过程的一系列踪迹. 这是与 Freeman 的感觉和创造性思维理论相关联的.

3.10.3　具延迟的神经网络的动力学问题及分析

在生物神经网络中，不同的神经元具有不同的触角长度和递质释放关系，这决定了神经元的兴奋传递时间必须是有差异的，以后延迟特性是其所固有的. 在人工神经网络中由于受硬件实现中开关滞后、参数的变化、分布杂散参数释放的有限 GB 特性等的影响，使延迟特性也是固有的，因此必须考虑具延迟的神经网络模型. 比较典型的是如下形式的具延迟的 Hopfield 模型

$$c_i\frac{\mathrm{d}u_i(t)}{\mathrm{d}t}=-\frac{u_i(t)}{R_i}+\sum_{j=1}^{N}J_{ij}f_j(u(t-\tau_i))\quad(i=1,2,\cdots,N).\quad(3.10.5)$$

一方面对于具延迟的人工神经网络模型，其中的滞后值 τ_j 一般情形下不太大. 很显然，如果 τ_j 很小，其退后模型的动力学行为与无退后模型的动力学行为是类似的. 但问题是要找到这个很小的 τ_j 的临界值 τ_c，否则应用时比较困难. 另一方面对于具延迟的生物神经网络的模型，实际情况延迟后的变化有一定的范围，必须讨论延迟后从零渐渐变大时的系统的动力学行为的变化.

1. 一维延迟模型的动力学问题

Parkdaman 和 Malta (1997, Neural Compute) 讨论了如下的具延迟的一维系统

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=-x(t)+A\tanh\,(x(t-\tau)),\qquad(3.10.6)$$

从理论上得到这样的结果，在参数的一定范围内，(3.10.6) 可能出现稳定收敛解或振荡周期解，但不会出现混沌奇异吸引子.

Pakdaman 和 Malta (1997, Neural Computa) 讨论了如下的具延滞的一维系统

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -x(t) + A\tanh\,(x(t-\tau)), \tag{3.10.7}$$

其理论结果是: 在参数的一定范围内, 该方程可能出现稳定收敛解或振荡周期解, 但不会出现混沌奇异吸引子.

Caianiello 和 Deluca 在 1966 年的文献 [41] 中讨论了以下形式的 Neuronic 方程 (也称决定方程)

$$u(t+\tau) = H\left[A(t) - \int_{-\infty}^{t} k(t-s)u(s)\mathrm{d}s - c\right],$$

这里 H 为单位阶跃函数

$$H(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leqslant 0, \end{cases}$$

其中 $u(t)$ 代表神经响应, 其值假定为 0 或 1, τ 表示系统的延滞, $A(t)$ 代表神经元的外部激励, $k(t)$ 表示神经元被激活或响应以后的耐受范围, 即从 0 到 1, c 代表神经元的阈值.

K.Gopalsamy 和 I.K.C.Leung 在文献 [42] 中对上述 Neuronic 方程适当变形所得到的下列方程的稳定行为作了讨论:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -x(t) + A\tanh\,(x(t)) + B\tanh\,(x(t-\tau)). \tag{3.10.8}$$

文献 [42] 中取 $A > 0, B < 0$. 相应的理论结果是: 在 A–B 参数平面中, 当系统参数位于由 $A = 0, B = 0$ 以及 $A - B = 1$ 所构成的三角区域中时, 系统状态变量对于任意的 τ 值都是全局渐近稳定的.

1998 年, Pakdaman 和 Malta 在文献 [43] 中考察了系统 (3.10.8) 的其它参数区域, 得到如下的理论结果:

在 A–B 参数平面中 (如图 3.10.6), 当 $A + B \leqslant 1$ 时, 0 点是唯一平衡点; 当 $A + B > 1$ 时, 有 3 个平衡点 $x_1 = a, x_2 = 0, x_3 = -a$, 这里 a 是满足 $a = (A+B)\tanh\,(a)$ 的唯一正实数.

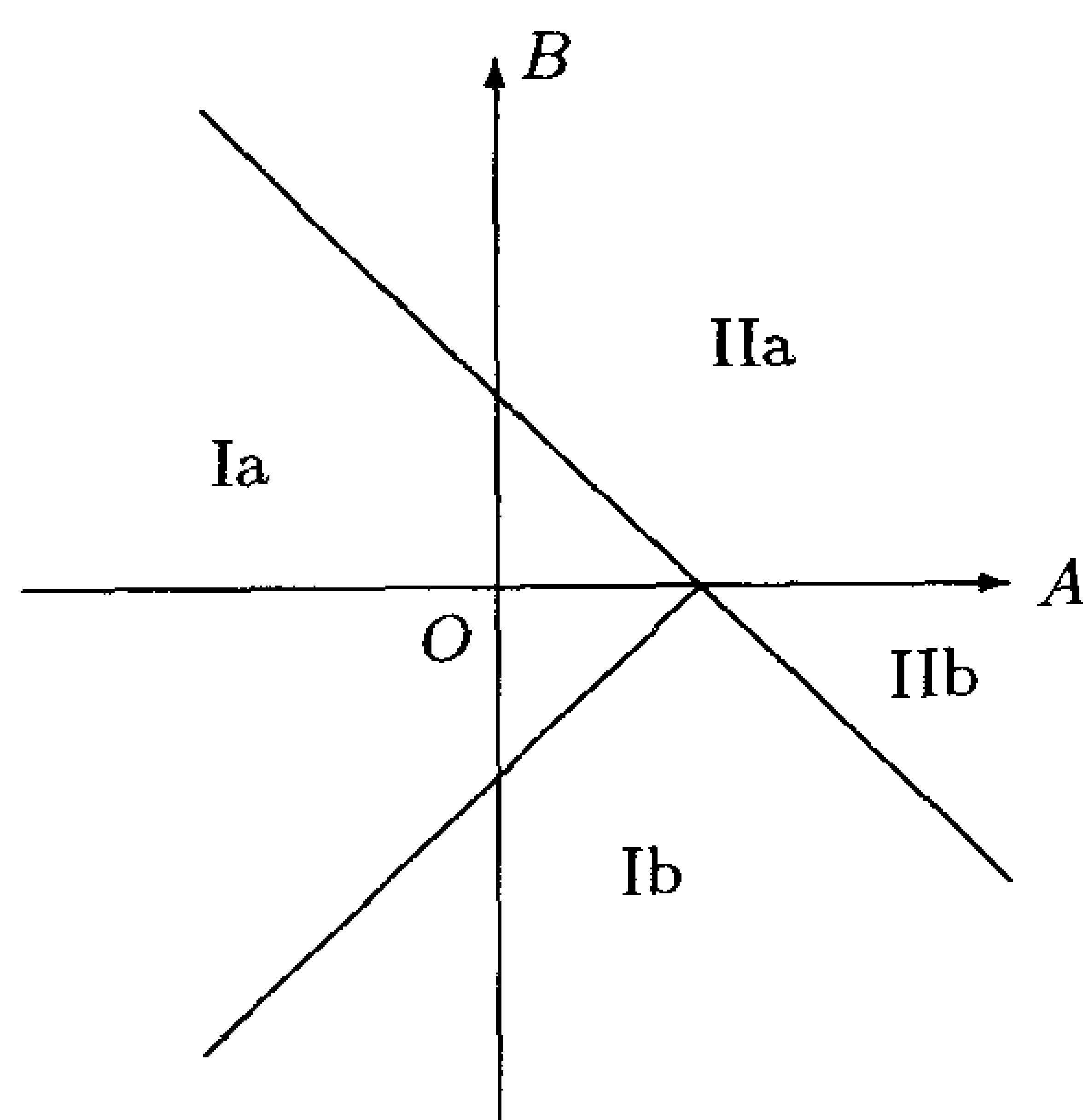


图 3.10.6

(1) 当 $A+B<1$, $A-B<1$ (即图 3.10.6 中的 Ia 区域) 时, 对任意的 $\tau\geqslant 0$, 0 点是全局渐近稳定的.

(2) 当 $A+B<1$, $A-B>1$ (即图 3.10.6 中的 Ib 区域) 时, 存在一个临界点 $\tau_c=\dfrac{\cos^{-1}\left(\dfrac{1-A}{B}\right)}{\sqrt{B^2-(1-A)^2}}$, 当 $0<\tau<\tau_c$ 时, 0 点是全局渐近稳定的; 当 $\tau\geqslant\tau_c$ 时, 原点在 τ_c 处产生 Hopf 分支进而失去稳定性, 即出现非衰减的振荡解, 有可能出现混沌现象.

(3) 当 $A+B>1$, $B>0$ (即图 3.10.6 中的 IIa 区域) 时, 对所有的 $\tau\geqslant 0$, $x_1=a$ 和 $x_3=-a$ 是局部渐近稳定的, 而 x_2 是不稳定的. 对任意的 τ, 大多数解趋于两个稳定平衡点中的某一个, 对于小的延滞, 当分离两个局部稳定域的边界恰好是原点的稳定流形时, 原系统是收敛的, 然而随着延滞的增大, 原点经历连续的 Hopf 分叉从而出现周期轨. 因而, 对于大的延滞, 存在着并不收敛到任意平衡点的解, 但在实际应用中, 振荡解不会出现, 系统的动力学行为是双稳定的收敛网络.

(4) 当 $A+B>1$, $B<0$ (即图 3.10.6 中的 IIb 区域) 时, 随着 τ 的增大, 产生不稳定现象, 这与 Ib 的情况类似, 不再详述. 需要说明的是: 对于小的延滞, 系统是收敛的, 但是随着延滞的增大, 两个稳定的平衡点通过产生周期振荡的 Hopf 分支点失去稳定

性时，收敛性消失. 因而对于大的延滞，大多数解表现为不衰减的周期振荡，有可能出现混沌现象.

上述的讨论结果仅限于定性描述，为此我们给出了严格的理论证明.

我们考虑如下自治方程:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(x_t),$$

其中 $f: C \to \mathrm{R}$ 全连续， $C = C[-\tau, 0]$, 且方程的解对初值连续依赖，过 $(0, \phi)$ 的解记为 $x(\phi)$.

假设 $V: C \to \mathrm{R}$ 是连续函数，定义 V 沿该方程的解的全导数为

$$\dot{V}(\phi) = \limsup_{h \to 0^+} \frac{1}{h}[V(x_h(\phi)) - V(x(\phi))].$$

对 C 中子集 G, 若 $V: C \to \mathrm{R}$ 在 G 的闭包 $\overline{G}$ 上连续，且在 G 上 $\dot{V} \leqslant 0$, 则称 V 是 G 上的 Lyapunov 泛函.

引理 3.10.1 设 $V: C \to \mathrm{R}$ 连续，$V(0) = 0$, 又存在非负函数 $\mu(r)$ 和 $\omega(r)$, 使得 $r \to \infty$ 时， $\mu(r) \to \infty$, 并且 $\mu(|\phi(0)|) \leqslant V(\phi)$, $\dot{V}(\phi) \leqslant -\omega(|\phi(0)|)$, 则方程 $\dfrac{\mathrm{d}x(t)}{\mathrm{d}t} = f(x_t)$ 的零解是稳定的，且每一个解都是有界的. 若 $\omega(r)$ 是正定的，则当 $t \to \infty$ 时，所有解趋于零.

定理 3.10.1 对于方程 (3.10.8), 当参数满足 $|B| < 1 - A$ (即参数平面中的 Ia 区域) 时，方程的零解是全局渐近稳定的.

证 首先由 $A + B < 1$ 保证了零解是唯一的平衡点.

取 Lyapunov 泛函

$$V(\phi) = -\frac{1}{A-1}\phi^2(0) + \int_{-\tau}^{0} \phi^2(\theta)\mathrm{d}\theta,$$

令 $\mu(r) = -\frac{1}{A-1}r^2$, 由于 $A<1$, 则

$$V(\phi) = -\frac{1}{A-1}\phi^2(0) + \int_{-\tau}^{0}\phi^2(\theta)\mathrm{d}\theta \geqslant -\frac{1}{A-1}\phi^2(0) = \mu(|\phi(0)|).$$

又

$$\begin{aligned}
\dot{V}|_{x_t} &= -\frac{2}{A-1}x(t)\dot{x}(t) + [x^2(t) - x^2(t-\tau)] \\
&= -\frac{2}{A-1}x(t)[-x(t) + A\tanh\,(x(t)) + B\tanh\,(x(t-\tau))] \\
&\quad +[x^2(t) - x^2(t-\tau)] \\
&= -\frac{2}{A-1}x(t)\Big[-x(t) + Ax(t) + Bx(t-\tau) \\
&\quad +A\sum_i \frac{\tanh^{(i)}(0)}{i!}x^i(t) + B\sum_j \frac{\tanh^{(j)}(0)}{j!}x^j(t)\Big] \\
&\quad +[x^2(t) - x^2(t-\tau)] \\
&= -\left[x(t) + \frac{B}{A-1}x(t-\tau)\right]^2 - \left(1 - \frac{B^2}{(A-1)^2}\right)x^2(t-\tau) \\
&\quad -\frac{2A}{A-1}O(x^3(t)) - \frac{2B}{A-1}O(x^3(t-\tau)) \\
&\leqslant -\left[x(t) + \frac{B}{A-1}x(t-\tau)\right]^2 - \frac{2A}{A-1}O(x^3(t)) \\
&\quad -\frac{2B}{A-1}O(x^3(t-\tau)) \\
&\leqslant -\frac{1}{2}\left[x(t) + \frac{B}{A-1}x(t-\tau)\right]^2 \\
&\leqslant -\frac{1}{2}\left[1 - \frac{B^2}{(A-1)^2}\right]x^2(t).
\end{aligned}$$

记 $\omega(r) = \frac{1}{2}\left[1 - \frac{B^2}{(A-1)^2}\right]r^2$, 由定理条件知, $\omega(r) > 0$, 即 $\omega(r)$ 是正定的. 从而由引理 3.10.1, 当 $t \to +\infty$ 时, $\lim\limits_{t\to+\infty} x(t) = 0$, 即零解是全局渐近稳定的, 证毕.

引理 3.10.2 (延滞方程的 Hopf 分支定理) 如果方程 (3.10.8)

满足如下条件:

[H1]　其对应的线性方程的特征方程具有一个纯虚根 $\lambda_0 = \mathrm{i}\,v_0 \neq 0$, 并且对任意的整数 m, 所有的特征根 $\lambda_j \neq \lambda_0, \overline{\lambda}_0$ 满足 $\lambda_j \neq m\lambda_0$;

[H2]　$\mathrm{Re}\,\lambda'(0) \neq 0$.

则方程 (3.10.8) 存在非常数周期解, 周期近似为 $\dfrac{2\pi}{v_0}$.

定理 3.10.2　对于方程 (3.10.8), 当参数满足 $A+B<1$, $A-B>1$ (即参数平面中的 Ib 区域) 时, 存在延滞的某一临界值 τ_c, 使得方程的零解在 τ_c 处发生 Hopf 分叉, 失去稳定性.

证　方程 (3.10.8) 对应的线性方程为

$$\frac{\mathrm{d}x}{\mathrm{d}t} = (A-1)x(t) + Bx(t-\tau), \tag{3.10.9}$$

其特征方程为 $\lambda = A - 1 + B\mathrm{e}^{-\lambda\tau}$.

当 $\tau = 0$ 时, $\lambda = A + B - 1$, 由 $A + B < 1$ 可知, $\tau = 0$ 处特征根具有负实部, 平衡点是渐近稳定的. 置 τ 从 0 增大, 欲寻找一个临界值 τ_c, 当 $\tau \in [0, \tau_c)$ 时, 特征根具有负实部, 而当 $\tau = \tau_c$ 时特征方程有纯虚根.

令 $\lambda = \mathrm{i}\,v$ $(v \neq 0)$, 代入特征方程得

$$\mathrm{i}\,v = A - 1 + B\mathrm{e}^{-\mathrm{i}\,v\tau} = (A - 1 + B\cos v\tau) - \mathrm{i}\,B\sin v\tau.$$

实虚部分别相等, 有

$$\begin{cases} A - 1 + B\cos v\tau = 0, \\ v + B\sin v\tau = 0. \end{cases}$$

由 $\sin v\tau^2 + \cos v\tau^2 = 1$ 得 $v = \sqrt{B^2 - (A-1)^2}$. 进而得

$$\tau_c = \frac{\cos^{-1}\left[\dfrac{1-A}{B}\right]}{\sqrt{B^2 - (A-1)^2}}.$$

现考察引理中的两个条件，显然 [H1] 满足．又对于特征方程 $\lambda - Be^{-\lambda\tau} - A + 1 = 0$，有 $\lambda'(\tau) + Be^{-\lambda\tau}(\tau\lambda'(\tau) + \lambda) = 0$，即 $\lambda'(\tau) + B\tau e^{-\lambda\tau}\lambda'(\tau) + \lambda Be^{-\lambda\tau} = 0$，从而

$$\lambda'(\tau) = \frac{-\lambda Be^{-\lambda\tau}}{1 + B\tau e^{-\lambda\tau}},$$

得到 $\lambda'(0) = -B\lambda(0)$，所以 $\mathrm{Re}\,\lambda'(0) = -B\mathrm{Re}\,\lambda(0) \neq 0$．可知 [H2] 也满足．

故由引理 3.10.2 知，方程 (3.10.8) 的零解在 τ_c 处产生 Hopf 分支，方程出现周期解．当稳定性丧失时，可以大致了解振荡周期的一些性质．在参量集 (A, B, τ_c) 处，方程 (3.10.9) 有一个解 $x(t) = e^{i v\tau}$，v 为角频率，由角频率和周期 T 的关系可知 $T = \dfrac{2\pi}{v}$，所以

$$T = \frac{2\pi}{\sqrt{B^2 - (A-1)^2}} = \frac{2\pi\tau_c}{\cos^{-1}\left[\dfrac{1-A}{B}\right]}.$$

由于 $1 - A$ 与 B 异号，容易求出 $2\tau \leqslant T \leqslant 4\tau$．这样可以界定方程 (3.10.8) 在不稳定点处振荡周期的范围．定理证毕．

以下讨论当参数满足 $A + B > 1$ (参数平面中的 IIa, IIb 区域)，即具有 3 个平衡点 $x_1 = a$, $x_2 = 0$, $x_3 = -a$ 时的情形．

引理 3.10.3 如果方程

$$\dot{x}(t) = ax(t) + bx(t-\tau) + \sum_{i+j\geqslant 2} c_{ij}(t)x^i(t)x^j(t-\tau) \tag{3.10.10}$$

满足如下条件：

(i) $a + b > 0$；

(ii) 存在 $\varepsilon > 0$，使得当 $|x| < \varepsilon$, $|y| < \varepsilon$ 时有 $\sum\limits_{i+j\geqslant 2} c_{ij}|x^i||y^j| < +\infty$．

则对任意 $\tau > 0$，方程 (3.10.10) 的零解是不稳定的，这里 $|c_{ij}(t)| \leqslant c_{ij}$．

定理 3.10.3 在 $A + B > 1$ 的情况下，方程 (3.10.8) 的零解

是不稳定的.

由引理 3.10.3 易得，证明从略.

现将方程 (3.10.8) 在 $x_1 = a$ 处 Taylor 展开，得到

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -x(t) + A\left[\frac{\mathrm{e}^a - \mathrm{e}^{-a}}{\mathrm{e}^a + \mathrm{e}^{-a}} + \frac{4}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}(x(t) - a) + \cdots\right]$$
$$+B\left[\frac{\mathrm{e}^a - \mathrm{e}^{-a}}{\mathrm{e}^a + \mathrm{e}^{-a}} + \frac{4}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}(x(t-\tau) - a) + \cdots\right].$$

令 $z(t) = x(t) - a$, 只考虑线性部分，则得到原方程的变分方程为

$$\frac{\mathrm{d}z}{\mathrm{d}t} = -z(t) + \frac{4A}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}z(t) + \frac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}z(t-\tau),$$

其相应的特征方程为

$$\lambda = -1 + \frac{4A}{(\mathrm{e}^a + \mathrm{e}^{-a})^2} + \frac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}\mathrm{e}^{-\lambda\tau}.$$

考察 $\tau = 0$ 时，特征方程变为

$$\lambda = -1 + \frac{4A}{(\mathrm{e}^a + \mathrm{e}^{-a})^2} + \frac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}.$$

在 $A+B>1$ 的前提下 (保证有 3 个平衡点), 取定 A, B 后，令 $f(a) = \mathrm{e}^{2a} - \mathrm{e}^{-2a} - 4a$, 由于 $f'(a) = 2\mathrm{e}^{2a} + 2\mathrm{e}^{-2a} - 4 = 2(\mathrm{e}^a - \mathrm{e}^{-a})^2 > 0$, 从而可知 $f(a) > f(0) = 0\ (a > 0)$, 即总有 $\dfrac{4(A+B)}{(\mathrm{e}^a + \mathrm{e}^{-a})^2} < 1$ 成立，即 $\tau = 0$ 时 $\lambda < 0$, 则在 $\tau = 0$ 处特征根具有负实部. 由 λ 对于 τ 的连续依赖性可知，对于充分小的 τ 值，特征根具有负实部，从而在 $A+B>1$ 时平衡点 $x_1 = a$ 是局部渐近稳定的. 特别地，在 $B<0$ 的情况下，置 τ 从 0 不断增大，欲寻找临界值 τ_c', 观察此处动力学行为的变化.

令 $\lambda = \mathrm{i}v$, 将其代入特征方程，有

$$\mathrm{i}v = -1 + \frac{4A}{(\mathrm{e}^a + \mathrm{e}^{-a})^2} + \frac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}\mathrm{e}^{-\mathrm{i}v\tau}$$

$$= -1 + \frac{4A}{(\mathrm{e}^a + \mathrm{e}^{-a})^2} + \frac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}(\cos v\tau - \mathrm{i}\sin v\tau)$$
$$= -1 + \frac{4A}{(\mathrm{e}^a + \mathrm{e}^{-a})^2} + \frac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}\cos v\tau - \mathrm{i}\,\frac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}\sin v\tau.$$

实虚部分别相等，有

$$\begin{cases} -1 + \dfrac{4A}{(\mathrm{e}^a + \mathrm{e}^{-a})^2} + \dfrac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}\cos v\tau = 0, \\ \\ v + \dfrac{4B}{(\mathrm{e}^a + \mathrm{e}^{-a})^2}\sin v\tau = 0. \end{cases}$$

由 $\sin v\tau^2 + \cos v\tau^2 = 1$ 得

$$v = \sqrt{\frac{16(B^2 - A^2) + 8A(\mathrm{e}^a + \mathrm{e}^{-a})^2}{(\mathrm{e}^a + \mathrm{e}^{-a})^4} - 1},$$

进而得

$$\tau_c' = \cos^{-1}\left[\frac{(\mathrm{e}^a + \mathrm{e}^{-a})^2 - 4A}{4B}\right]/v.$$

类似于定理 3.10.2, 在 τ_c' 处，平衡点 $x_1 = a$ 失稳，产生 Hopf 分支. 对 $x_3 = -a$ 的讨论情况类似. 因而对于具有 3 个平衡点的情形， $x_2 = 0$ 是不稳定的， $x_1 = a, x_3 = -a$ 是局部渐稳的.

定理 3.10.4 方程 (3.10.8) 的所有解是有界的.

证 由方程 (3.10.8) 可得

$$\begin{aligned} \left|\frac{\mathrm{d}x(t)}{\mathrm{d}t} + x(t)\right| &= |A\tanh\,(x(t)) + B\tanh\,(x(t-\tau))| \\ &\leqslant |A||\tanh\,(x(t))| + |B||\tanh\,(x(t-\tau))| \\ &\leqslant |A| + |B|, \end{aligned}$$

故

$$-(|A| + |B|) \leqslant \frac{\mathrm{d}x(t)}{\mathrm{d}t} + x(t) \leqslant |A| + |B|.$$

对于 $\dfrac{\mathrm{d}x(t)}{\mathrm{d}t}+x(t)\leqslant|A|+|B|$, 两边同乘 e^t 然后从 t_0 到 t 积分得到

$$x\mathrm{e}^t-x_0\mathrm{e}^{t_0}\leqslant\int_{t_0}^{t}(|A|+|B|)\mathrm{e}^s\mathrm{d}s,$$

从而

$$x\leqslant(|A|+|B|)(1-\mathrm{e}^{t_0-t})+x_0\mathrm{e}^{t_0}\mathrm{e}^{-t}.$$

所以

$$\begin{cases}x\leqslant|A|+|B|+x_0\mathrm{e}^{t_0} & (x_0>0),\\ x<|A|+|B| & (x_0<0),\end{cases}$$

右有界得证.

同理可得

$$\begin{cases}x>-(|A|+|B|) & (x_0>0),\\ x\geqslant-(|A|+|B|)+x_0\mathrm{e}^{t_0} & (x_0<0),\end{cases}$$

左有界得证.

综上所述,

$$\begin{cases}-(|A|+|B|)<x\leqslant|A|+|B|+x_0\mathrm{e}^{t_0} & (x_0>0),\\ -(|A|+|B|)+x_0\mathrm{e}^{t_0}\leqslant x<|A|+|B| & (x_0<0),\end{cases}$$

所以 $x(t)$ 有界，证毕.

综上所述, 由于方程 (3.10.8) 的所有解是有界的, 因而在系统的平衡点不稳定的情形下, 方程 (3.10.8) 的解可能会在一个有限的范围内荡来荡去, 轨道的动力学行为非常复杂, 并非完全随机却又永不重复，即出现混沌动力学行为. 图 3.10.7 正是这些情形的数

值模拟图.

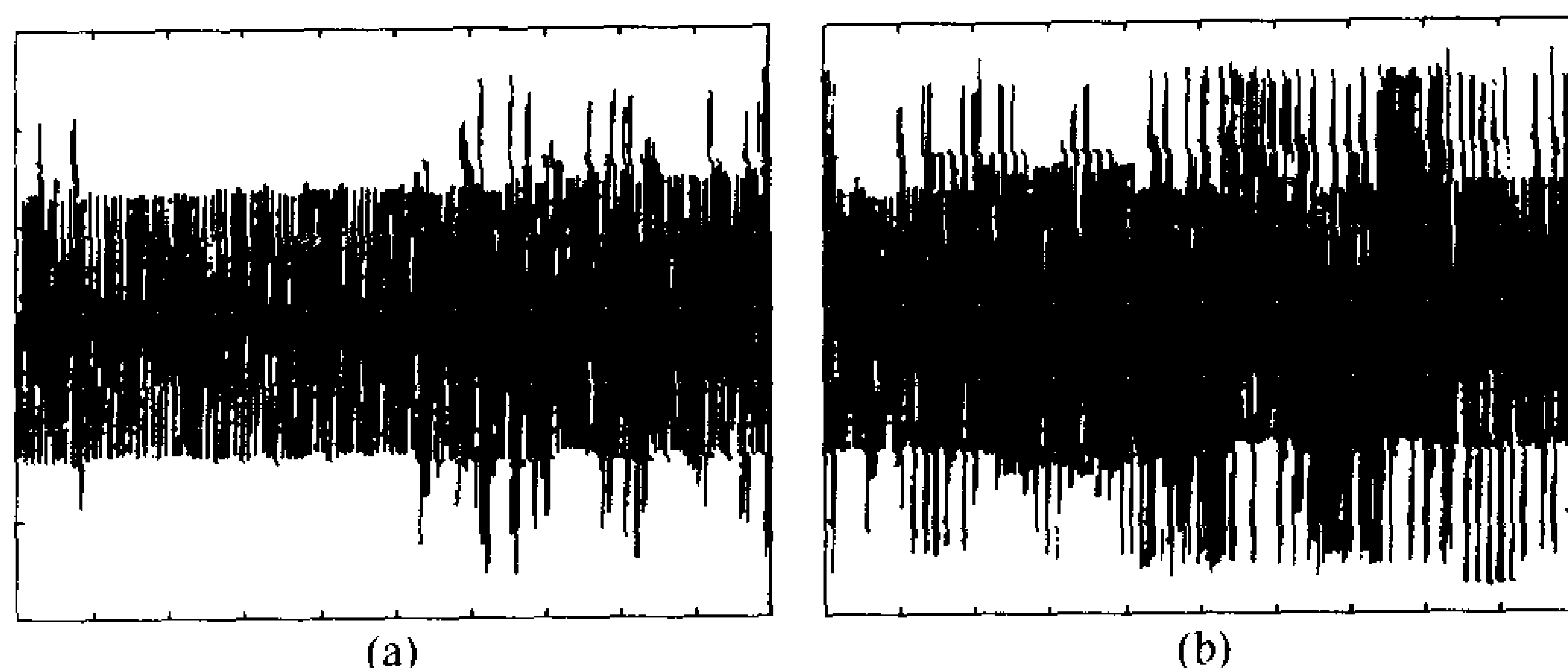


图 3.10.7　取参数 $A=13.5373$, $B=-11.4627$, $\tau_c'=2.009$, 初始条件为 $\phi(\theta)=5-5\theta^2/\tau^2$, $\theta\in[0,\tau]$. (a) $\tau=3.0$, Lyapunov 指数 $L=1.2531$; (b) $\tau=3.9$, Lyapunov 指数 $L=1.4618$

2. 二维延迟模型的动力学问题

(1) 二维模型

Babcock 和 Westerrelt (1987, Phys. D.) 讨论了具延迟的二维系统

$$\begin{cases} \dfrac{\mathrm{d}u_1(t)}{\mathrm{d}t}=A_2\tanh\ (u_2(t-\tau))-u_1(t), \\ \dfrac{\mathrm{d}u_2(t)}{\mathrm{d}t}=A_1\tanh\ (u_1(t-\tau))-u_2(t). \end{cases} \tag{3.10.11}$$

他们从数值计算的角度得到下面的结果, 随着 A_1, A_2 及 τ 的不同选取, 系统 (3.10.11) 可能有不稳定的不动点及不稳定的极限环, 但是没有得到二维系统可能出现复杂的混沌动态的结果.

阮炯等人 (1994) 讨论了如下的一般的二维神经系统:

$$\begin{cases} \dfrac{\mathrm{d}u_1(t)}{\mathrm{d}t}=-R_1u_1(t)+T_{11}\tanh\ (b_1u_1(t-\tau)) \\ \qquad\qquad +T_{12}\tanh\ (b_2u_2(t-\tau)), \\ \dfrac{\mathrm{d}u_2(t)}{\mathrm{d}t}=-R_2u_2(t)+T_{21}\tanh\ (b_3u_1(t-\tau)) \\ \qquad\qquad +T_{22}\tanh\ (b_4u_2(t-\tau)), \end{cases} \tag{3.10.12}$$

其中 $\tau > 0$, R_1, R_2, b_k $(k = 1,2,3,4)$, T_{ij} $(i,j = 1,2)$ 分别是可选取的参数，在神经网络模型及非线性电路方程中均有各自的物理意义.

显然， (3.10.12) 存在平凡的吸引子 $(0,0)$, 并且一般地还可能存在其它非平凡的吸引子 (u_1^0, u_2^0), $(u_1^0)^2 + (u_2^0)^2 \neq 0$. 又容易验证若存在非平凡的吸引子 (u_1^0, u_2^0), 则 $(-u_1^0, -u_2^0)$ 也是系统的非平凡的吸引子. 关于系统 (3.10.12) 的吸引子具体的个数讨论很复杂，但有一点可以肯定， (3.10.12) 只能有奇数个吸引子.

现在讨论 (3.10.12) 较一般的系统

$$\frac{\mathrm{d}u_i(t)}{\mathrm{d}t} = -R_i u_i(t) + \sum_{j=1}^{2} T_{ij} f(b_{ij} u_j(t-\tau_i)) + I_i, \quad i = 1,2 \tag{3.10.13}$$

在参数的不同取值范围内解的动力学性质. 这里参数 R_i, b_{ij}, $\tau_i > 0$, T_{ij}, I_i $(i,j = 1,2)$ 均有各自的物理意义，一般设 $I_i \neq 0$, $f(x) = \tanh(x)$, $t \in [t_0, +\infty)$. 初始条件为

$$u_i(t) = \phi_i(t), \quad t \in [-\tau, 0], \quad \tau = \max_{i=1,2}\{\tau_i\}, \quad \phi_i(t) \in C([-\tau, 0], \mathrm{R}). \tag{3.10.14}$$

注意 (3.10.13) 当 $I_i \neq 0$ 时，平衡点的存在及个数的问题不是很容易解决的.

(2) 理论结果

引理 3.10.4 系统 (3.10.13) 满足初始条件 (3.10.14) 的解在 $[0, +\infty)$ 存在且有界.

证 解的局部存在性可由分步法得到，解在 $[0, +\infty)$ 上的存在性可由下面的有界性并用延拓定理得到. 设 $(u_1(t), u_2(t))$ 是 (3.10.13) 满足 (3.10.14) 的解，则

$$-R_i u_i(t) - \sum_{j=1}^{2} |T_{ij}| - |I_i| \leqslant \dot{u}_i(t) \leqslant -R_i u_i(t) + \sum_{j=1}^{2} |T_{ij}| + |I_i|, \quad i = 1,2.$$

由微分不等式知对于任意 $t>0$, $u_i(t)$ 有界. 证毕.

引理 3.10.5 若系统 (3.10.13) 中参数满足下列的条件

$$\frac{1}{R_i}\sum_{j=1}^{2}|T_{ij}|b_{ij}\leqslant\eta\quad(i=1,2),\tag{3.10.15}$$

其中 $0<\eta<1$, 则系统 (3.10.13) 至少存在一个平衡点.

证 若 (u_1^*,u_2^*) 是 (3.10.13) 的平衡点, 则满足

$$u_i^*=R_i^{-1}\sum_{j=1}^{2}T_{ij}f(b_{ij}u_j^*)+R_i^{-1}I_i\quad(i=1,2).$$

写成向量形式为

$$u^*=g(u^*)+q\triangleq F(u^*),$$

这里

$$\begin{aligned}u^*&=(u_1^*,u_2^*)^{\mathrm{T}},\\ g(u^*)&=\left(R_1^{-1}\sum_{j=1}^{2}T_{1j}f(b_{1j}u_j^*),R_2^{-1}\sum_{j=1}^{2}T_{2j}f(b_{2j}u_j^*)\right)^{\mathrm{T}},\\ q&=(R_1^{-1}I_1,R_2^{-1}I_2)^{\mathrm{T}},\end{aligned}$$

则 u^* 是 $F:\mathrm{R}^2\to\mathrm{R}^2$ 的不动点. 因为

$$\|g(u^{(1)})-g(u^{(2)})\|\leqslant R_i^{-1}\sum_{j=1}^{2}|T_{ij}|b_{ij}\|u^{(1)}-u^{(2)}\|,$$

所以在条件 (3.10.15) 成立时, 有

$$\|g(u^{(1)})-g(u^{(2)})\|\leqslant\eta\|u^{(1)}-u^{(2)}\|,\quad 0<\eta<1.$$

故 F 为压缩映射, 即 F 必有不动点 u^* 存在. 证毕.

定理 3.10.5 若在系统 (3.10.13) 中参数满足条件 (3.10.15) 以及

$$R_i^{-1}\sum_{j=1}^{2}|T_{ji}|b_{ji}<1,\tag{3.10.16}$$

则 (3.10.13) 有唯一的平衡点且是全局渐近稳定的.

证 由引理 3.10.2 知系统 (3.10.13) 的平衡点存在, 下证唯一性. 若存在两个不同的平衡点 $u^{(1)}, u^{(2)}$, 则成立下式

$$\|u^{(1)} - u^{(2)}\| = \|g(u^{(1)}) - g(u^{(2)})\| \leqslant R_i^{-1} \sum_{j=1}^{2} |T_{ij}| b_{ij} \|u^{(1)} - u^{(2)}\|,$$

即 $R_i^{-1} \sum_{j=1}^{2} |T_{ij}| b_{ij} \geqslant 1$, 与条件 (3.10.15) 矛盾, 故 (3.10.13) 仅有唯一的平衡点.

以下由 Lyapunov 给出的方法证明该唯一的平衡点是全局渐近稳定的. 先设 (u_1^*, u_2^*) 为 (3.10.13) 的唯一的平衡点, 令 $y_i(t) = u_i(t) - u_i^*$ $(i = 1, 2)$, 则 (3.10.13) 可以化成

$$\frac{\mathrm{d}y_i(t)}{\mathrm{d}t} = -R_i y_i(t) + \sum_{j=1}^{2} T_{ij} f(b_{ij}(y_j(t - \tau_j) + u_j^*)), \quad i = 1, 2. \tag{3.10.17}$$

定义 Lyapunov 泛函

$$V(\phi) = \frac{1}{2} \sum_{i=1}^{2} \phi_i^2(0) + \frac{1}{2} \sum_{i=1}^{2} \int_{t-\tau_i}^{t} \phi_i^2(s) \mathrm{d}s,$$
$$\phi_i \in C([-\tau, 0], \mathrm{R}), \quad i = 1, 2,$$

则 V 沿方程组 (3.10.17) 的全导数为

$$\begin{aligned}
&\left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{(3.10.17)} \\
&= \sum_{i=1}^{2} y_i(t) \frac{\mathrm{d}y_i(t)}{\mathrm{d}t} + \frac{1}{2} \sum_{i=1}^{2} R_i [y_i^2(t) - y_i^2(t - \tau_i)] \\
&= \sum_{i=1}^{2} y_i(t) \left[-R_i y_i(t) + \sum_{i=1}^{2} T_{ij} b_{ij} f'(\theta_{ij}(t)) y_j(t - \tau_j) \right] \\
&\quad + \frac{1}{2} \sum_{i=1}^{2} R_i y_i^2(t) - \frac{1}{2} \sum_{i=1}^{2} R_i y_i^2(t - \tau_i)
\end{aligned}$$

$$\leqslant -\frac{1}{2}\sum_{i=1}^{2}R_i y_i^2(t)-\frac{1}{2}\sum_{j=1}^{2}R_j y_j^2(t-\tau_j)+\frac{1}{2}\sum_{i,j=1}^{2}|T_{ij}|b_{ij}y_i^2(t)$$

$$+\frac{1}{2}\sum_{i,j=1}^{2}|T_{ij}|b_{ij}y_j^2(t-\tau_j)$$

$$=-\frac{1}{2}\sum_{i=1}^{2}\left[R_i-\sum_{j=1}^{2}|T_{ij}|b_{ij}\right]y_i^2(t)$$

$$-\frac{1}{2}\sum_{i=1}^{2}\left[R_i-\sum_{j=1}^{2}|T_{ij}|b_{ij}\right]y_i^2(t-\tau_i)$$

$$\leqslant -\frac{1}{2}\sum_{i=1}^{2}\left[R_i-\sum_{j=1}^{2}|T_{ij}|b_{ij}\right]\|y_i^2(t)\|$$

$$\leqslant -\frac{1}{2}\eta\|y_i^2(t)\|,$$

其中用到了 (3.10.15) 与 (3.10.16) 两个条件. 又 $\theta_{ij}(t)$ 是位于 $y_j(t-\tau_j)$ 与 $y_j(t-\tau_j)+u_j^*$ 之间的值. 所以由 Lyapunov 泛函的基本方法得知 (u_1^*,u_2^*) 是 (3.10.13) 的全局渐近稳定的平衡点. 证毕.

以下假定 $R_1=R_2=R$, $\tau_1=\tau_2=\tau$, $I_1=I_2=0$, 则 (3.10.13) 变为

$$\frac{\mathrm{d}u_i(t)}{\mathrm{d}t}=-R_iu_i(t)+\sum_{j=1}^{2}T_{ij}f(b_{ij}u_j(t-\tau_j))\quad(i=1,2).\quad(3.10.18)$$

将 (3.10.18) 在平衡点 (0,0) 附近展开得

$$\begin{cases}\dfrac{\mathrm{d}u_1(t)}{\mathrm{d}t}=-Ru_1(t)+T_{11}b_{11}u_1(t-\tau)+T_{12}b_{12}u_2(t-\tau),\\[2mm]\dfrac{\mathrm{d}u_2(t)}{\mathrm{d}t}=-Ru_2(t)+T_{21}b_{21}u_1(t-\tau)+T_{22}b_{22}u_2(t-\tau),\end{cases}\quad(3.10.19)$$

它对应的特征方程为

$$\det((\lambda+R)\mathrm{e}^{\tau\lambda}E-\overline{T}),$$

其中 $\overline{T} = (T_{ij}b_{ij})$, E 为单位阵. 记 $\sigma(\overline{T})$ 为 $\overline{T}$ 的谱, $S(\overline{T}) = \max\{\text{Re } u : u \in \sigma(\overline{T})\}$, $\rho(\overline{T}) = \max\{|u| : u \in \sigma(\overline{T})\}$, 则有

$$(\lambda + R)e^{\tau\lambda} = u, \quad u \in \sigma(\overline{T}). \tag{3.10.20}$$

令 $u = \alpha + i\beta$, $\alpha, \beta \in \text{R}$, $\lambda = iv$ 代入 (3.10.20), 分开实部与虚部后, 有

$$\begin{cases} R\cos\tau v - v\sin\tau v = \alpha, \\ v\cos\tau v + R\sin\tau v = \beta, \end{cases}$$

即

$$\begin{pmatrix} R & -v \\ v & R \end{pmatrix} \begin{pmatrix} \cos\tau v \\ \sin\tau v \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

解出 $\cos\tau v$, $\sin\tau v$, 得

$$\cos\tau v = \frac{R\alpha + v\beta}{R^2 + v^2}, \quad \sin\tau v = \frac{R\beta - v\alpha}{R^2 + v^2}, \tag{3.10.21}$$

因而有

$$(R\alpha + v\beta)^2 + (R\beta - v\alpha)^2 = (R^2 + v^2)^2,$$

即

$$v^4 + (2R^2 - |u|^2)v^2 + R^2(R^2 - |u|^2) = 0. \tag{3.10.22}$$

定理 3.10.6 若 (3.10.15) 条件满足, 且 $\rho(\overline{T}) < R$, 则对任意时滞 τ_i, $(0,0)$ 是 (3.10.18) 的渐近稳定的平衡点, 而且是唯一的平衡点.

证 首先由 $S(\overline{T}) < \rho(\overline{T})$, 当 $\rho(\overline{T}) < R$ 时, $S(\overline{T}) - R < 0$, 即 (3.10.19) 的平衡点 $(0,0)$ 是渐近稳定的.

若 $\tau < 0$, 由 $\rho(\overline{T}) < R$, 任意 $u \in \sigma(\overline{T})$, 则关于 v^2 的二次方程 (3.10.22) 仅有实部为负的解, 即关于 v 无实根. 所以不论 $\tau > 0$ 如何选取, 关于 λ 的超越方程 (3.10.20) 都没有纯虚根, 即 (3.10.20) 的根的实部永远为负, 从而 $(0,0)$ 是 (3.10.19) 的渐近稳定的平衡

点. 由泛函微分方程的一次近似理论知 (0,0) 是 (3.10.18) 的渐近稳定平衡点, 且这结论不依赖于时滞 τ_i. 证毕.

定理 3.10.7 若 (3.10.15) 条件满足, 且 $S(\overline{T}) < R < \rho(\overline{T})$, 显然要求 T 为非对称. 则存在 $\tau_{c_0} > 0$, 当 $\tau \in [0, \tau_{c_0})$ 时, (0,0) 是系统 (3.10.18) 的渐近稳定的平衡点; 当 $\tau \geqslant \tau_{c_0}$ 时, (0,0) 失稳; 随着时滞 τ 的增大, 对每个 τ 值 (3.10.18) 可能出现周期解或混沌解或周期解与混沌解并存的 3 种情况.

证 由 $S(\overline{T}) < R$, 当 $\tau = 0$ 时 (3.10.18) 的平衡点 (0,0) 是渐近稳定的. 由 (3.10.20) 知, 当 $\rho(\overline{T}) > R$ 时, 存在 $u \in \sigma(\overline{T})$ 使得 $R^2 - |u|^2 < 0$, 因而关于 v^2 的二次方程 (3.10.22) 有正实根, 关于 v 的四次方程有根, 记为 v_c, 且可以由 (3.10.21) 知,

$$\tau_{ck} = \frac{1}{v_c} \arccos \frac{R\alpha + \beta v_c}{R^2 + v_c^2} + \frac{1}{v_c} 2k\pi, \quad k = 0, 1, 2, \cdots, \tag{3.10.23}$$

即 $\tau = \tau_{ck}$ $(k = 0, 1, 2, \cdots)$ 时, 关于 λ 的方程 (3.10.20) 有纯虚根. 所以当 $\tau \in [0, \tau_{c_0})$ 时, (3.10.20) 的特征根实部全为负数, 即此时 (0,0) 是 (3.10.19) 的渐近稳定的平衡点, 从而也是 (3.10.18) 的渐近稳定的平衡点.

在 (3.10.20) 中视 λ 为 τ 的函数, 可以求得

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\tau} = -\frac{\lambda(\lambda + R)}{1 + (\lambda + R)\tau}, \tag{3.10.24}$$

经计算得

$$\mathrm{Re} \left. \frac{\mathrm{d}\lambda}{\mathrm{d}\tau} \right|_{\tau = \tau_{ck}} = v_c^2 > 0. \tag{3.10.25}$$

易知 (3.10.19) 在 τ_{ck} 处出现 Hopf 分支, 从而 (3.10.18) 出现周期解, 随着 τ 的增大还可能出现混沌或周期解相间的情况. 具体要通过计算 Lyapunov 指数来判别是否出现混沌.

定理 3.10.8 若 (3.10.15) 条件满足, 且 $R < S(\overline{T})$, (0,0) 是系统 (3.10.18) 的不稳定的平衡点, (3.10.18) 的解不可能收敛于某

个平衡点，只可能出现周期解、混沌等各种复杂动力学行为.

证 由 $R < S(\overline{T})$, 当 $\tau = 0$ 时 (3.10.20) 有正实部的特征根，所以 $(0,0)$ 是 (3.10.19) 的不稳定的平衡点. 由 $R < \rho(\overline{T})$ 类似于定理 3.10.7 的证明求得 v_c, τ_{ck} $(k = 0, 1, 2, \cdots)$. 在 $\tau = \tau_{ck}$ 处，(3.10.20) 有纯虚根且 $\left.\dfrac{\mathrm{d}\lambda}{\mathrm{d}\tau}\right|_{\tau=\tau_{ck}} > 0$, 即 $(0,0)$ 是不稳定的，故可出现各种复杂的动态. 证毕.

下面讨论比较复杂的二维延迟网络系统

$$\begin{cases} u_1'(t) = -R_1 u_1(t) + T_{11}\tanh\,(b_1 u_1(t-\tau)) \\ \qquad\qquad +T_{12}\tanh\,(b_2 u_2(t-\tau)), \\ u_2'(t) = -R_2 u_2(t) + T_{21}\tanh\,(b_3 u_1(t-\tau)) \\ \qquad\qquad +T_{22}\tanh\,(b_4 u_2(t-\tau)), \end{cases} \tag{3.10.26}$$

其中 $\tau > 0$, $R_1 \neq R_2$, $R_i > 0$ $(i = 1, 2)$, T_{ij} $(i, j = 1, 2)$. 现对平凡的吸引子 $(0,0)$ 的稳定性态与时滞 τ 的关系进行讨论. (3.10.26) 关于 $(0,0)$ 的线性化方程为

$$\begin{cases} \overline{u}_1'(t) = -R_1\overline{u}_1(t) + T_{11}b_1\overline{u}_1(t-\tau) + T_{12}b_2\overline{u}_2(t-\tau), \\ \overline{u}_2'(t) = -R_2\overline{u}_2(t) + T_{21}b_3\overline{u}_1(t-\tau) + T_{22}b_4\overline{u}_2(t-\tau). \end{cases} \tag{3.10.27}$$

对应的特征方程为

$$\begin{aligned} H(\lambda) &= \lambda^2 + (R_1 + R_2)\lambda + R_1R_2 - (T_{11}b_1 + T_{22}b_4)\mathrm{e}^{-\lambda\tau}\lambda \\ &\quad -(R_1T_{22}b_4 + R_2T_{11}b_1)\mathrm{e}^{-\lambda\tau} \\ &\quad +(T_{11}T_{22}b_1b_4 - T_{12}T_{21}b_2b_3)\mathrm{e}^{-2\lambda\tau} \\ &= 0, \end{aligned} \tag{3.10.28}$$

当 $\tau = 0$ 时,

$$\begin{aligned} H(\lambda) &= \lambda^2 + (R_1 + R_2)\lambda - (T_{11}b_1 + T_{22}b_4)\lambda + R_1R_2 \\ &\quad +(T_{11}T_{22}b_1b_4 - T_{12}T_{21}b_2b_3 - R_1T_{22}b_4 - R_2T_{11}b_1) \\ &= 0, \end{aligned}$$

当下面的条件

$$\begin{cases} R_1+R_2-T_{11}b_1-T_{22}b_4>0, \\ R_1R_2+(T_{11}T_{22}b_1b_4-T_{12}T_{21}b_2b_3)-(R_1T_{22}b_4+R_2T_{11}b_1)>0 \end{cases} \tag{3.10.29}$$

满足时, (3.10.28) 的根具备负实部. 置 τ 从 0 增大, 依据特征方程根对参数 τ 的连续性, 下面寻找一个临界值 τ_c, 它具有如下性质: 当 $\tau\in[0,\tau_c)$ 时, (3.10.28) 的特征根均具负实部, 当 $\tau=\tau_c$ 时有纯虚根, 而当 $\tau>\tau_c$ 时, (3.10.28) 有正实部的特征根. 也就是说, 随 τ 从 0 增大到 τ_c, 然后增大到大于 τ_c, 平凡吸引子 (0,0) 的渐近稳定性发生变化. 这里用到微分差分方程组的稳定性的一次近似的结果, 能用方程组 (3.10.27) 的 (0,0) 稳定性结果去讨论方程组 (3.10.26) 的 (0,0) 稳定性.

下面求时滞的临界值 τ_c, 将 $\lambda=\mathrm{i}\,v_c$ 代入 (3.10.28), 分开实部与虚部得到下面两式

$$-v^2+R_1R_2-Av_c\sin v_c\tau+Bv_c\sin v_c\tau+c\cos 2\tau v_c=0, \quad (3.10.30)_1$$

$$(R_1+R_2)v_c-Av_c\cos v_c\tau-Bv_c\sin v_c\tau+c\sin 2\tau v_c=0, \quad (3.10.30)_2$$

其中 $A=T_{11}b_1+T_{22}b_4$, $B=R_1T_{22}b_4+R_2T_{11}b_1$, $C=T_{11}T_{22}b_1b_4-T_{12}T_{21}b_2b_3$. 将 $(3.10.29)_1\times\cos v_c\tau-(3.10.29)_2\times\sin v_c\tau$ 和 $(3.10.29)_1\times\sin v_c\tau+(3.10.29)_2\times\cos v_c\tau$ 写成矩阵向量的等式形式

$$\begin{pmatrix} -v_c^2+R_1R_2+c & -(R_1+R_2)v_c \\ (R_1+R_2)v_c & -v_c^2+R_1R_2-c \end{pmatrix}\begin{pmatrix} \cos v_c\tau \\ \sin v_c\tau \end{pmatrix}=\begin{pmatrix} B \\ Av_c \end{pmatrix}.$$

记

$$\begin{aligned} \Delta &= v_c^4+(R_1^2+R_2^2)v_c^2+(R_1^2+R_2^2-c^2), \\ \Delta_1 &= [A(R_1+R_2)-B]v_c^2+B(R_1R_2-c), \\ \Delta_2 &= v_c\{-Av_c^2+[A(R_1R_2^2+c)-B(R_1+R_2)]\}, \end{aligned}$$

易知有

$$\cos v_c\tau = \frac{\Delta_1}{\Delta}, \quad \sin v_c\tau = \frac{\Delta_2}{\Delta}, \quad \Delta_1^2 + \Delta_2^2 = \Delta^2.$$

定义

$$H(v) \triangleq \Delta_1^2(v) + \Delta_2^2(v) - \Delta^2(v) = 0,$$

注意它是 v 的最高 8 次的仅有偶数次幂的多项式, 因为 $H(+\infty) < 0$, 若下式成立

$$(R_1T_{22}b_4 + R_2T_{11}b_1)^2 - (R_1R_2 + T_{11}T_{22}b_1b_4 - T_{12}T_{21}b_2b_3)^2 > 0.$$

则有 $H(0) > 0$, 此时

$$H(0) = B^2(R_1R_2 - c)^2 - (R_1^2R_2^2 - c^2)^2 = (R_1R_2 - c)^2[B^2 - (R_1R_2 + c)^2].$$

所以存在 $v_c > 0$, 使得 $H(v_c) = 0$. 当然还有其它的一些存在 v_c 的充分条件, 一般可以直接求 8 次多项式 $H(v)$ 的正实根 v_c, 再从 $\cos(v_c\tau_c) = \dfrac{\Delta_1(v_c)}{\Delta(v_c)}$, 可以求得临界值 τ_c 为

$$\tau_c = \frac{1}{v_c}\cos^{-1}\left(\frac{\Delta_1(v_c)}{\Delta(v_c)}\right) + \frac{1}{v_c}2k\pi, \quad k = 0, 1, 2, \cdots. \tag{3.10.31}$$

注意, 当 $\tau \in [0, \tau_c)$ 时, $(0,0)$ 渐稳, 当 $\tau > \tau_c$ 时, $(0,0)$ 失稳. 类似地可以假定 $\tau = 0$ 时 (3.10.28) 有正实部的根, 让 τ 增大, 以找到时滞临界值 $\overline{\tau}$, 使得 $0 \leqslant \tau < \overline{\tau}$ 时, (3.10.28) 有正实部的根, 当 $\tau > \overline{\tau}$ 时, (3.10.28) 均有负实部的根. 换言之, $(0,0)$ 的不稳定性随 τ 的增大而变为渐稳. 假如满足前面的条件 (3.10.15), 则系统只有唯一的平衡点 $(0,0)$, 从而在 $\tau = \tau_c$ 处是 $(0,0)$ 稳定性发生改变的临界点, 从 (3.10.31) τ_c 的表达式可见, 随 k 取不同的自然数的值有一系列的 $\tau_c^{(k)}$, 从而可见有一系列的延迟的值 $\tau_c^{(k)}$, 使得 $(0,0)$ 的稳定性发生变化, 如果在 $\tau_c^{(k)}$ 处满足一些条件可以使之存在 Hopf 分支, 从而产生新的周期解, 又因为整个网络系统的有界性, 所以也就可能在平衡点 $(0,0)$ 不稳定情形下有混沌动力学行为出现, 也有可能周期振荡与混沌奇异吸引子并存.

当然如果满足下面的条件 (3.10.32), 则 $\tau = 0$ 时, $H(\lambda) = 0$ 有正实部的根, 从而 (0,0) 为 $\tau = 0$ 时的不稳定平衡点.

$$p < 0 \quad 或者 \quad p \geqslant 0 \quad 且 \quad q < 0, \tag{3.10.32}$$

其中 $p = R_1 + R_2 - T_{11}b_1 - T_{22}b_4$, $q = R_1R_2 + (T_{11}T_{22}b_1b_4 - T_{12}T_{21}b_2b_3) - (R_1T_{22}b_4 + R_2T_{11}b_1)$. 可以按照条件 (3.10.29) 满足时讨论类似地进行.

如果系统还存在非原点的平衡点, 则可以建立系统位于非原点平衡点的变分方程, 然后仿照前面的讨论, 可以得到类似的结果. 需要指出的是对存在非原点的平衡点情形, 就引出多个平衡点的稳定性的讨论, 从而其动力学行为变得很复杂, 这尚属没有完全解决的问题.

(3) 数值模拟

例 3.10.1 在 (3.10.26) 系统中选取参数值如下:

$$R_1 = 0.9, \quad R_2 = 0.73, \quad T_{11} = 0.5, \quad T_{12} = 0.7, \quad T_{21} = -25,$$

$$T_{22} = -0.7, \quad b_1 = 2.31, \quad b_2 = 1.91, \quad b_3 = 1.5, \quad b_4 = 1.71.$$

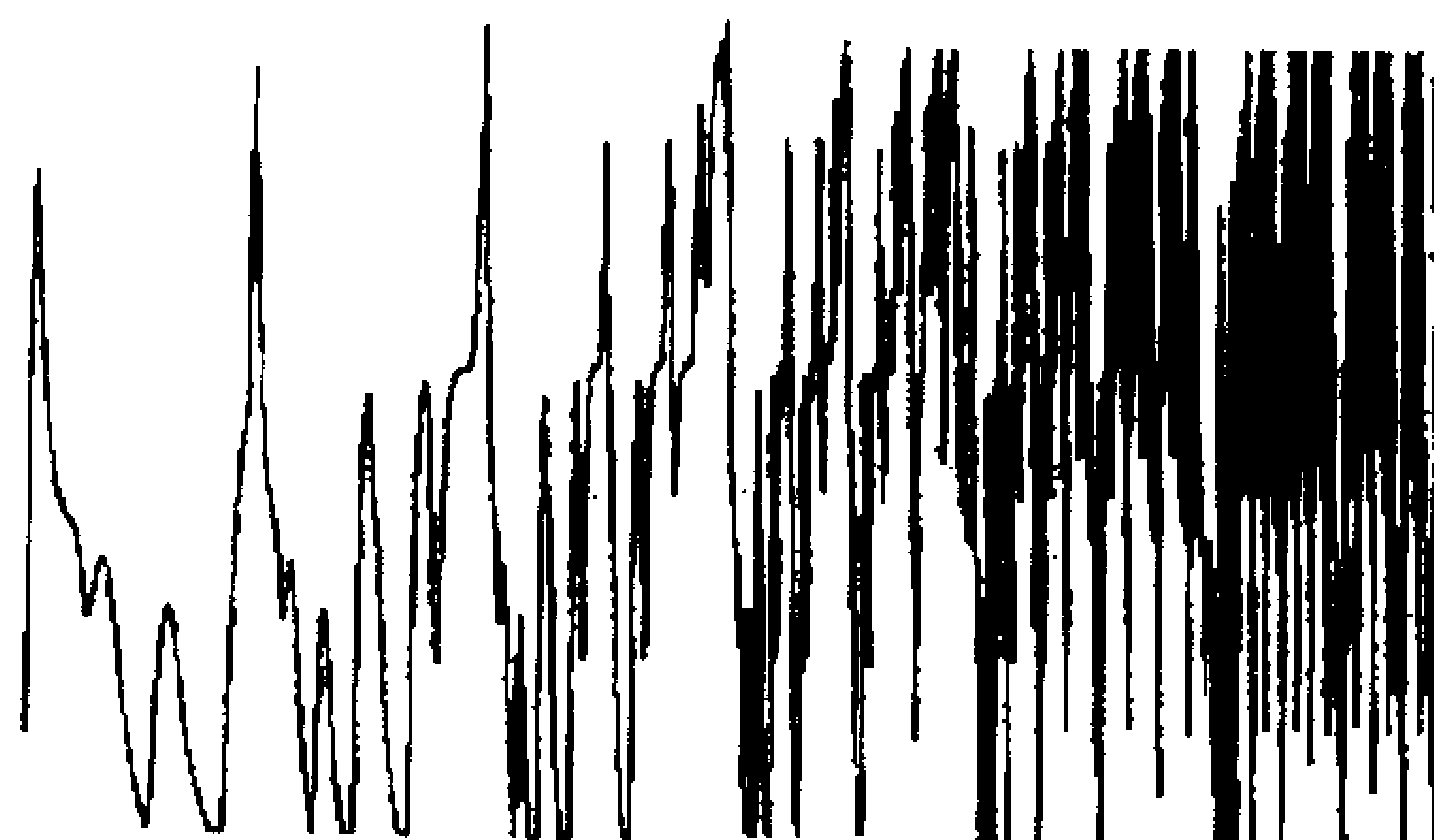
先看 $\tau = 0$ 的情形, 此时特征方程 (3.10.28) 的特征根均具负实部, 从而知 (0,0) 是系统 (3.10.26) 的渐进稳定的吸引子. 此时初始条件取值如下: 当 $\theta \in [0,\tau]$ 时, $u_i(\theta) = \phi_i(\theta)$ $(i = 1,2)$, $\phi_1(\theta) = \phi_2(\theta) = 5 - 5\dfrac{\theta}{\tau}$, τ 取为 1.7. 图 3.10.8(a) 表示 $u_1(t)$ 的解曲线图, 图 3.10.8(b) 表示 u_1, u_2 的相平面图.

如何选 τ 使得出现复杂动态, 可以把 τ 从 0 起按一定的步长选取, 看每个选定的 τ 值对应的 u_1, u_2 相平面图. 计算相应的 Lyapunov 指数, 开始时呈现非周期奇异吸引子.

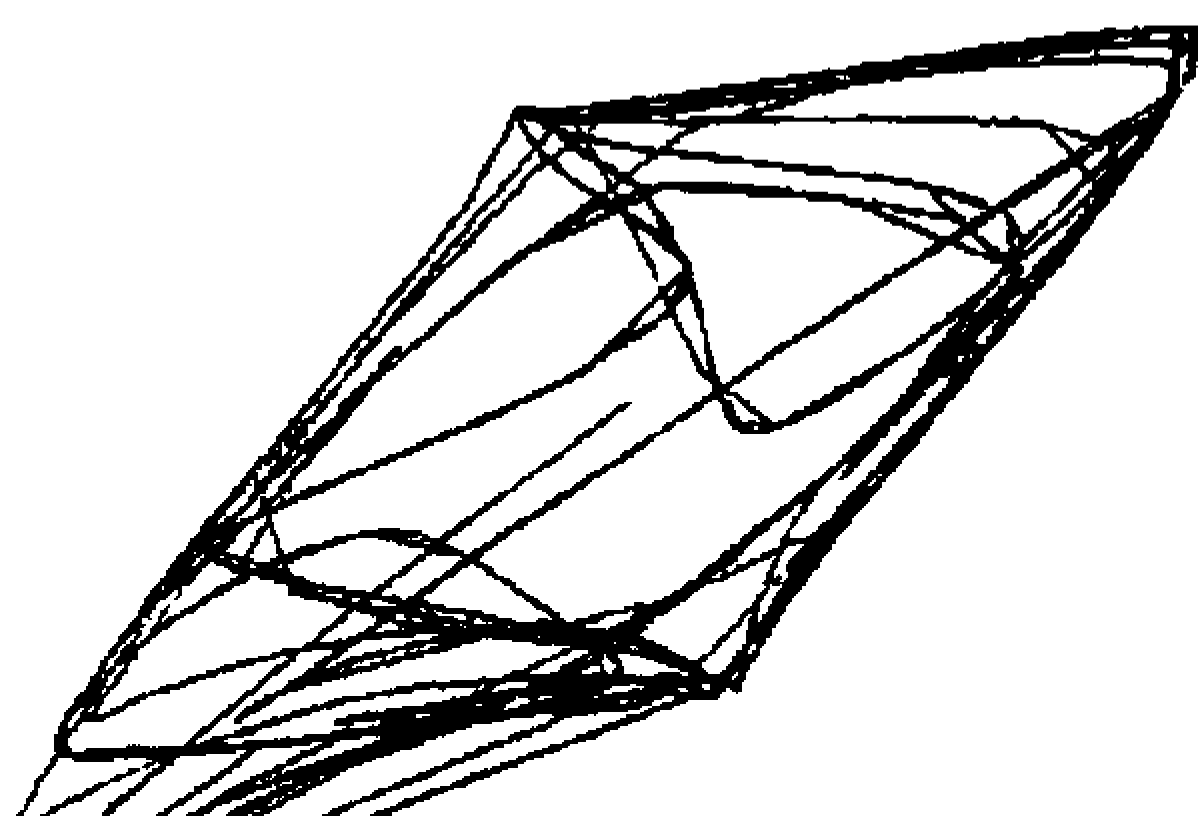
例 3.10.2 在系统 (3.10.26) 中选取参数如下:

$$R_1 = 3.1, \quad R_2 = 11.1, \quad T_{11} = 2.0, \quad T_{12} = -1.0, \quad T_{21} = 3.51,$$

$$T_{22} = 0.7, \quad b_1 = 2.31, \quad b_2 = 2.91, \quad b_3 = 2.1, \quad b_4 = 2.71.$$



(a)

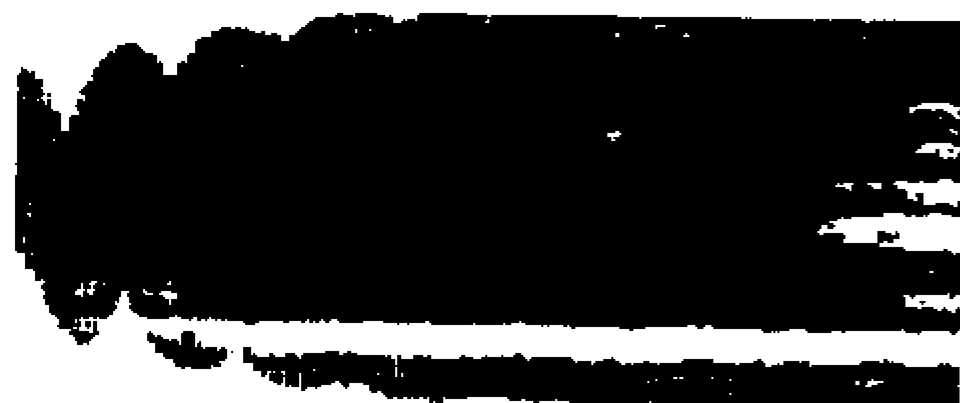


(b)

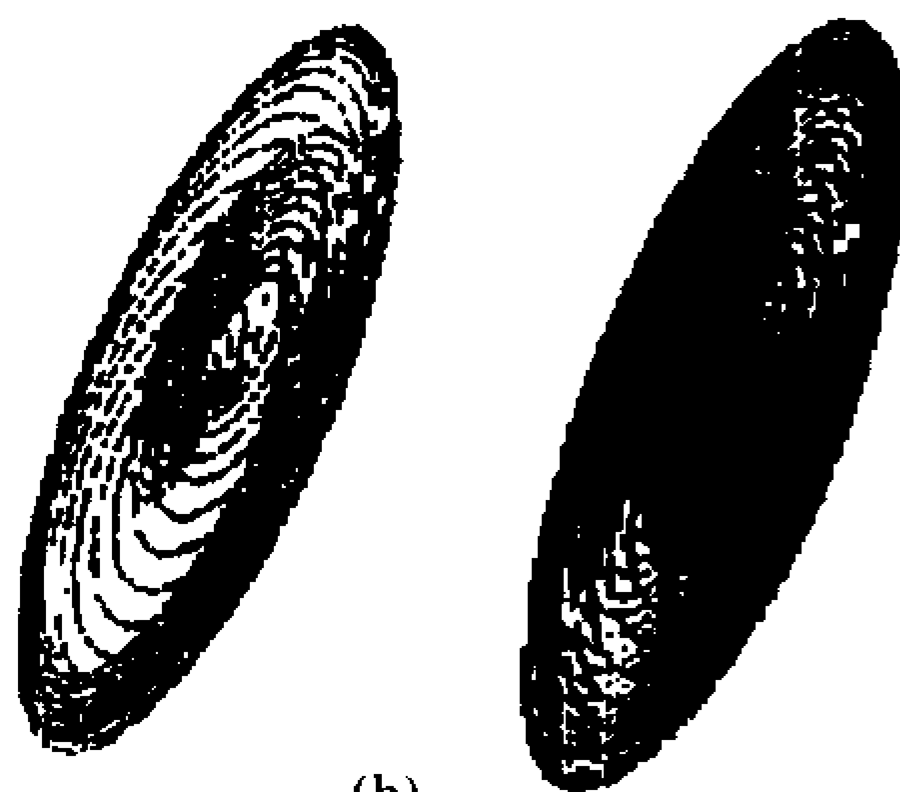
图 3.10.8

(a) $u_1(t)$ 解曲线图; (b) u_1, u_2 相平面图

先置 $\tau=0$, 由相应的特征方程的根均具负实部知该系统在无时滞时, 平凡吸引子 $(0,0)$ 为渐近稳定. 然后让 τ 增大, 按一定的步长, 在数值计算方面取 $\tau=5.81$, 初始条件为 $\theta\in[0,\tau]$ 时, $u_i(\theta)=\phi_i(\theta)$ $(i=1,2)$, $\phi_1(\theta)=0.45-0.45\dfrac{\theta}{\tau}$, $\phi_2(\theta)=0.5-0.5\dfrac{\theta}{\tau}$. 图 3.10.9(a) 表示 $u_1(t)$ 的解曲线图, 图 3.10.9(b) 表示 u_1, u_2 的相图.



(a)



(b)

图 3.10.9

(a) $u_1(t)$ 解曲线图; (b) $u_1(t)$, $u_2(t)$ 的相图

3. 高维 ($N>3$) 情形的延迟神经网络动力学

对于三维的延迟神经网络动力学方程组，考虑如下简化的形式

$$u_i'(t) = -u_i(t) + \sum_{j=1}^{3} T_{ij} f(b_{ij} u_i(t-\tau)) \quad (i=1,2,3), \qquad (3.10.33)$$

$(0,0,0)$ 是它的平凡吸引子，对应的特征方程如果设 $T_{ii}=0$ $(i=1,2,3)$ 就有如下形式:

$$H(\lambda) = 3\mathrm{e}^{3\lambda\tau}(\lambda+1)^3 + \mathrm{e}^{\lambda\tau}(\lambda+1)s_1 + s_0 = 0, \qquad (3.10.34)$$

其中

$$\begin{aligned} s_1 &= T_{31}T_{12}b_{31}b_{12} - T_{23}T_{31}b_{13}b_{32} - T_{12}T_{21}b_{12}b_{21}, \\ s_0 &= T_{32}T_{31}T_{13}b_{31}b_{13}b_{32} - T_{21}T_{13}T_{32}b_{21}b_{13}b_{32}. \end{aligned}$$

令 $z=(\lambda+1)\mathrm{e}^{\lambda\tau}$, 则 (3.10.34) 可化为

$$z^3 + s_1 z + s_0 = 0. \qquad (3.10.35)$$

设上述三次方程的 3 个根为 z_j $(j=1,2,3)$, 则 (3.10.34) 可以化为如下形式的 3 个方程

$$(\lambda+1)\mathrm{e}^{\lambda\tau} = z_j \quad (j=1,2,3). \qquad (3.10.36)$$

易见在二维情形这类方程已经有过讨论. 为了寻找 (3.10.36) 的纯虚根 $\lambda=\mathrm{i}v$ 存在的条件，将 (3.10.36) 的实部与虚部分开后，得到

$$\begin{aligned} v\cos v\tau + \sin v\tau &= \mathrm{Im}\ z_j, \\ \cos v\tau - v\sin v\tau &= \mathrm{Re}\ z_j, \end{aligned}$$

其中 $\mathrm{Re}\ z_j$ 及 $\mathrm{Im}\ z_j$ 分别为 z_j 的实部及虚部. 从这两式得

$$\cos v\tau = \frac{-v\mathrm{Im}\ z_j - \mathrm{Re}\ z_j}{-(v^2+1)}, \quad \sin v\tau = \frac{v\mathrm{Re}\ z_j - \mathrm{Im}\ z_j}{-(v^2+1)}.$$

为简单起见，记 $R = \text{Re } z_j$, $I = \text{Im } z_j$, 则有

$$(vI + R)^2 + (vR - I)^2 = (v^2 + 1)^2,$$

化简得 $(v^2+1)(v^2+1-R^2-I^2) = 0$. 该方程存在实解 v_0 的条件为 $R^2 + I^2 \geqslant 1$, 也就是要求三次方程 (3.10.35) 的根均不在单位圆内. 将此化为关于 s_0, s_1 的条件，记 $\Delta = \dfrac{s_0^2}{4} + \dfrac{s_1^3}{27}$, 分 3 种情形:

(1) $\Delta < 0$. 此时 (3.10.35) 有 3 个实根，其模为 $\dfrac{2\sqrt{|s_1|}}{\sqrt{3}}$, v_0 存在的条件为 $\dfrac{2\sqrt{|s_1|}}{\sqrt{3}} \geqslant 1$. 由于此时 $s_1 < 0$, 故条件化为 $s_1 \leqslant -\dfrac{3}{4}$;

(2) $\Delta = 0$. 此时 (3.10.35) 有一对重实根，模为 $\dfrac{2\sqrt[3]{|s_0|}}{\sqrt[3]{3}}$, 还有一个实根，其模为 $2\dfrac{2\sqrt[3]{|s_0|}}{\sqrt[3]{3}}$. 故 v_0 存在的条件为 $\dfrac{2\sqrt[3]{|s_1|}}{\sqrt[3]{3}} \geqslant 1$, 即 $|s_0| \geqslant 2$;

(3) $\Delta > 0$. 置 $\gamma_1 = \sqrt[3]{-\dfrac{s_0}{2} + \sqrt{\Delta}}$, $\gamma_2 = \sqrt[3]{-\dfrac{s_0}{2} - \sqrt{\Delta}}$, 此时 (3.10.35) 有一个实根为 $\gamma_1 + \gamma_2$, 还有一对共轭复根，其模为 $\sqrt{\gamma_1^2 + \gamma_2^2 - \gamma_1\gamma_2}$. 故 v_0 存在的条件为 $|\gamma_1 + \gamma_2| \geqslant 1$ 及 $\sqrt{\gamma_1^2 + \gamma_2^2 - \gamma_1\gamma_2} \geqslant 1$.

同理非平凡吸引子 (u_1^0, u_2^0, u_3^0) 的变分方程之特征方程的实根随 τ 的值从 0 变大到临界值，再由一个临界值到旁边一个临界值之间动力学行为可以类似地讨论. Marcus 和 Wang, L. 等 (1991, Phys. D., p.51) 给出了具体的三维系统

$$\begin{cases} \dot{u}_1(t) = -u_1(t) + 2.0\tanh\,(6.1u_2(t)) - 3.5\tanh\,(2.5u_3(t)), \\ \dot{u}_2(t) = -u_2(t) + 3.8\tanh\,(8.0u_1(t-\tau)), \\ \dot{u}_3(t) = -u_3(t) + 3.8\tanh\,(8.0u_1(t-\tau)) \end{cases} \tag{3.10.37}$$

的数值计算例子，他们的计算表明当 $\tau = 0.64\text{ms}$ 时系统 (3.10.37) 有周期解，当 $\tau = 0.97\text{ms}$ 时，系统 (3.10.37) 出现混沌，当 $\tau > 0.97\text{ms}$ 时，周期解和混沌都可能出现.

对于一般的 n 维延迟神经网络类型

$$c_i \frac{\mathrm{d}u_i(t)}{\mathrm{d}t} = -\frac{u_i(t)}{R_i} + \sum_{j=1}^{N} J_{ij} f_j(u(t-\tau_j)) \quad (i = 1, 2, \cdots, N), \tag{3.10.38}$$

当 $c_i = c$, $f_i = f$, $\tau_i = \tau$, $R_i = R$ 时，可以通过变换，将上式化为一次近似形式

$$\frac{\mathrm{d}u_i(t)}{\mathrm{d}t} = -u_i(t) + \sum_{j=1}^{N} \beta J_{ij} u_j(t-\tau) \quad (i = 1, 2, \cdots, N). \tag{3.10.39}$$

若 J 的特征值为 N 个实根 $\lambda_1, \lambda_2, \cdots, \lambda_N$, $\lambda_{\max}$ 与 $\lambda_{\min}$ 分别表示最大特征值和最小特征值. 要使上面的假设能成立，需要 $J = J^{\mathrm{T}}$ 的条件，通过未知函数的代换可以将 (3.10.39) 化为如下形式:

$$\frac{\mathrm{d}x_i(t)}{\mathrm{d}t} = -x_i(t) + \beta\lambda_i x_i(t-\tau), \quad i = 1, 2, \cdots, N. \tag{3.10.40}$$

利用特征方程的分析方法可以得出如下结论:

(1) 当 $\beta < \dfrac{1}{\lambda_{\max}}$, $\dfrac{1}{\beta} > -\lambda_{\min}(\omega^2+1)^{1/2}$, $\omega = -\tan(\omega\tau)$, $\dfrac{\pi}{2} < \omega\tau < \pi$ 时，原点是稳定的;

(2) 当 $\beta > \dfrac{1}{\lambda_{\max}}$ 或者 $\beta < \dfrac{1}{\lambda_{\max}}$ 且 $\dfrac{1}{\beta} < -\lambda_{\min}(\omega^2+1)^{1/2}$, $\omega = -\tan(\omega\tau)$, $\dfrac{\pi}{2} < \omega\tau < \pi$ 时，原点是不稳定的，从而 $\beta = \dfrac{1}{\lambda_{\max}}$ 是叉形分支点，而在 $\beta = -\lambda_{\min}(\omega^2+1)^{1/2}$, $\omega = -\tan(\omega\tau)$, $\dfrac{\pi}{2} < \omega\tau < \pi$ 处发生 Hopf 分支.

如果记 $\|J\|_2$ 表示 $J = (J_{ij})$ 的 2 范数 ($J^{\mathrm{T}}J$ 的最大特征值的平方根)，现在讨论如下形式的方程组:

$$\frac{\mathrm{d}u_i(t)}{\mathrm{d}t} = -u_i(t) + \sum_{j=1}^{N} J_{ij} f(\beta_j u_j(t-\tau_j)), \tag{3.10.41}$$

这里 f 的形式可取 S 型函数, 例如 tanh, 记 $\beta=\max\{\beta_1,\cdots,\beta_N\}$.

定理 3.10.9 若 $\beta\|J\|_2<1$, 则系统 (3.10.41) 只有一个全局渐近稳定的平衡点.

证 首先证明平衡点是唯一的, 即只有原点是平衡点. 设系统 (3.10.41) 有非平凡的不动点 $x^0=(x_1^0,x_2^0,\cdots,x_N^0)^{\mathrm{T}}$, 则

$$x_i^0=\sum_{j=1}^{N}J_{ij}f_j(x_j^0),\quad i=1,2,\cdots,N.$$

令 $y^0=(y_1^0,y_2^0,\cdots,y_N^0)^{\mathrm{T}}$, 其中 $y_i^0=f_i(x_i^0)$, $i=1,2,\cdots,N$, 则 $x^0=Jy^0$, 因此 $(x^0)^{\mathrm{T}}x^0=(x^0)^{\mathrm{T}}Jy^0$, 于是

$$\|x^0\|^2=\|x^0\|_2^2\leqslant\|x^0\|\|J\|_2\|y^0\|,$$

又 $\|y_j^0\|=|f_j(y_j^0)|\leqslant\beta_j|x_j^0|\leqslant\beta|x_j^0|$, 所以

$$\|y^0\|\leqslant\beta\|x^0\|,\quad \|x^0\|\leqslant\beta\|J\|_2\|x^0\|^2,$$

于是 $\beta\|J\|_2\geqslant1$, 与已知条件矛盾, 故只能 $\|x^0\|=0$, 即原点是系统 (3.10.41) 的唯一平衡点.

以下证明原点是渐近稳定的. 设 $\phi=(\phi_1,\phi_2,\cdots,\phi_N)^{\mathrm{T}}$, 取 V 泛函

$$V(\phi)=\sum_{i=1}^{N}\phi_i^2(0)+\sum_{i=1}^{N}\int_{-\tau_i}^{0}\phi_i^2(\theta)\mathrm{d}\theta,$$

以及 $a(r)=r^2$, 则当 $r\to+\infty$ 时, $a(r)\to+\infty$, 显然 $a(|\phi(0)|)\leqslant V(\phi)$.

V 通过 (3.10.41) 的解的全导数为

$$\begin{aligned}&\left.\frac{\mathrm{d}V(u)}{\mathrm{d}t}\right|_{(3.10.41)}\\&=2\sum_{i=1}^{N}u_i(t)\dot{u}_i(t)+\sum_{i=1}^{N}u_i^2(t)-\sum_{i=1}^{N}u_i(t-\tau_i)\\&=-\sum_{i=1}^{N}u_i^2(t)-\sum_{i=1}^{N}u_i^2(t-\tau_i)+2\sum_{i=1}^{N}\sum_{j=1}^{N}u_i(t)J_{ij}f_j(u_j(t-\tau_j)).\end{aligned}$$

记 $U(t)=(u_1(t),u_2(t),\cdots,u_N(t))^{\mathrm{T}}$, $\eta=(f_1(u_1(t-\tau_1)),f_2(u_2(t-\tau_1)),\cdots,f_N(u_N(t-\tau_1)))^{\mathrm{T}}$. 所以成立

$$\|\eta\|\leqslant\beta\sqrt{\sum_{i=1}^N u_i^2(t-\tau_1)},\quad \|U(t)\|^2=\sum_{i=1}^N u_i^2(t),$$

$$U^{\mathrm{T}}(t)J\eta\leqslant\|U(t)\|\|J\|_2\beta\left[\sum_{i=1}^N u_i^2(t-\tau_i)\right]^{1/2},$$

从而

$$\begin{aligned}&\left.\frac{\mathrm{d}V(u)}{\mathrm{d}t}\right|_{(3.10.41)}\\ \leqslant&-\sum_{i=1}^N u_i^2(t)-\sum_{i=1}^N u_i^2(t-\tau_i)+2\sum_{i=1}^N\sum_{j=1}^N u_i(t)J_{ij}f_j(u_j(t-\tau_j))\\ \leqslant&-(1-\beta\|J\|_2)\|U(t)\|^2+(\beta\|J\|_2-1)\sum_{i=1}^N u_i^2(t-\tau_i)\\ \leqslant&-(1-\beta\|J\|_2)\|U(t)\|^2.\end{aligned}$$

若定义 $b(r)=(1-\beta\|J\|_2)|r|^2$, 易知, 当 $1-\beta\|J\|_2\geqslant 0$ 时, $b(r)$ 正定. 所以原点是系统的渐近稳定平衡点. 证毕.

本定理的一个例子, 设

$$J=\frac{1}{2}\begin{pmatrix}0&-1&-1\\-1&0&-1\\-1&-1&0\end{pmatrix},$$

其 3 个特征值为 $\lambda_1=\lambda_2=\dfrac{1}{2}$, $\lambda_3=-1$. 由本定理知, 当 $\beta<1$ 时, 对任何滞后, 原点是唯一的全局渐近稳定的平衡点. 又按分支理论, $\dfrac{1}{2}$ 是偶重的特征根, 故 $\beta=2$ 处不可能发生分支. -1 是奇重特征根, 具有出现分支的可能, 当然还存在不少问题需要研究.

关于一般的具有时滞的神经网络系统，当时滞很小时，即接近于无时滞的情形，它的稳定性与零时滞的网络系统情形的稳定性是一样的. 问题是如何给出时滞的界限 τ^*, 即当 $\tau \in [0,\tau^*)$ 时，有时滞与零时滞两个网络系统在稳定性方面是等价的.

一般地，考虑如下系统

$$C\frac{\mathrm{d}x(t)}{\mathrm{d}t} = -Q^{-1}x(t) + Ay(x) + Bz(x), \tag{3.10.42}$$

其中 $x,y,z \in \mathrm{R}^n$, $C,Q,A,B \in \mathrm{R}^{n\times n}$ 为常矩阵,

$$\begin{aligned}
C &= \mathrm{diag}(C_1,\cdots,C_n), \quad C_i > 0, i=1,2,\cdots,n,\\
Q &= \mathrm{diag}(Q_1,\cdots,Q_n), \quad Q_i > 0, i=1,2,\cdots,n,\\
y(x) &= (f_1(x_1(t)),\cdots,f_n(x_n(t)))^{\mathrm{T}},\\
z(x) &= (f_1(x_1(t-\tau_1)),\cdots,f_n(x_n(t-\tau_n)))^{\mathrm{T}},
\end{aligned}$$

其中 $\tau_i \geqslant 0$ 是时滞， $f_i(s)$ $(s\in \mathrm{R})$ 是 sigmoid 函数，并满足

$$0 < f_i(s) \leqslant \beta_i^{-1}, \quad s \in \mathrm{R}.$$

$$\lim_{s\to-\infty} f_i(s) = m_i^- > -\infty, \quad \lim_{s\to+\infty} f_i(s) = m_i^+ > +\infty.$$

假定 $g_i = f_i^{-1}$, 则

$$\dot{g}_i(s) \geqslant \beta_i, \quad s \in (m_i^-, m_i^+)$$

且 $G(y) = \mathrm{diag}(\dot{g}_1(y_1),\cdots,\dot{g}_n(y_n))$ 是正定矩阵.

如果 $A+B$ 对称，则系统 (3.10.42) 称为对称的，对这种情形有如下定理.

定理 3.10.10 如果存在 $p>0$, 使得

$$\frac{1}{2}\|B\|(p+p^{-1}\tau_i^2) < C_i\beta_i, \quad i=1,2,\cdots,n, \tag{3.10.43}$$

其中 $\|B\|=\sup\dfrac{\|Bx\|}{\|x\|}$, $\|x\|=\sqrt{\sum\limits_{i=1}^{n}x_i^2}$, 那么 (3.10.42) 的所有解都趋向于该系统的平衡点集.

对于 A, B 满足 $A=DA^*$, $B=DB^*$ (或 $A=A^*D$, $B=B^*D$), $D=\mathrm{diag}(d_1,\cdots,d_n)$, $d_i>0$, $i=1,2,\cdots,n$, $T^*=A^*+B^*$ 是对称的条件情形，有以下结果.

定理 3.10.11 如果存在 $p>0$, 使得

$$\frac{1}{2}\|B^*\|d_i(p+p^{-1}\tau_i^2)<C_i\beta_i,\quad i=1,2,\cdots,n,\tag{3.10.44}$$

那么 (3.10.42) 的所有解都趋向于该系统的平衡点集.

定理的证明比较复杂, 这里省略了. 由于 (3.10.43) 及 (3.10.44) 条件不大容易验证，所以下面提供两个推论供应用时使用.

推论 3.10.1 在定理 3.10.11 对 A, B 假设的条件下，显然定理 3.10.10 的 A, B 条件是定理 3.10.11 的 A, B 条件的特例. 如果

$$\mu\|B^*\|d_i\tau_i<C_i\beta_i,\quad i=1,2,\cdots,n,\tag{3.10.45}$$

其中 $\mu=\frac{1}{2}(\sqrt{\eta\xi^{-1}}+\sqrt{\xi\eta^{-1}})$, $\xi=\max\{\tau_i,i=1,2,\cdots,n\}$, $\eta=\min\{\tau_i,i=1,2,\cdots,n\}$. (如果最简单的情形 $\tau_i=\tau$, $i=1,2,\cdots,n$, 则 $\mu=1$, 从而 (3.10.43) 变为 $\tau<\min\limits_i\left[\dfrac{C_i\beta_i}{d_i\|B^*\|}\right]$.) 那么系统 (3.10.40) 的所有解都趋向于该系统的平衡点集.

推论 3.10.2 在定理 3.10.11 对 A, B 的假设条件下，如果

$$\|B^*\|d_i\tau<C_i\beta_i,\quad i=1,2,\cdots,n,\tag{3.10.46}$$

其中 $\tau=\max\{\tau_i,i=1,2,\cdots,n\}$, 那么 (3.10.42) 的所有解都趋向

于该系统的平衡点集.

实际上, 下面的条件与 (3.10.44) 是等价的,

$$\tau < \min_i \left[\frac{C_i \beta_i}{d_i \|B^*\|} \right] = \frac{1}{\|B^*\|} \min_i \left[\frac{C_i \beta_i}{d_i} \right]. \tag{3.10.47}$$

因此, 实际上给出了 τ_i 的界限 τ^*, 使得在 $\tau_i \in [0, \tau_i^*)$ 时, 系统的所有解都趋向于系统的平衡点集的性质成立.

本节的定理 3.10.1~3.10.11 以及推论 3.10.1~3.10.2 的详细细节可参见下面所引的参考文献 [10~17].

参 考 文 献

[1] S. A. Elias & S. Grossberg, Pattern formation, contrast control, and oscillation in the short term memory of shunting on–center off–surround networks, Biol. Cybern, 20(1975), 69~98.

[2] W. J. Freeman & G. Viana Di Prisco, EEG Spetial pattern differences with discriminated odors manifest chaotic an limit cycle attractors in olfactory bulb of rabbits, In Brain Theory, G. Palm (ed.), New York: Springer-Verlag, 1986.

[3] W. J. Freeman & G. Viana Di Prisco, Correlation of olfactory EEG with behavior, Time series analysis, Behavioral Neuroscience, 100(1986b), 753~763.

[4] L. B. Almeida, A learning rule for asynchronous perceptrons with feedback in a combinatorial environment, Proceedings of IEEE International Conference on Neural Networks, Vol.2, Amari, S. A. (1982): Competitive and cooperative aspects in dunamics of neural excitation and self–organization, In: Competition and Cooperation in Neural Nets, S. Amari and M. Arbib eds., Springer Lecture Notes in Biomath, 45.

[5] R. J. Williams & D. Zipser, A learning algorithm for continually

running fully recurrent neural networks, Technical report ICS 8805, Inst. Cognitive Science, U. California at San Diego.

[6] D. Serran–Schréber & A. Cleeremans, Learning and internal representations using analog CMOS processing arrays, In: Advanced Research in VLSI: Proceedings of the 1987 Stanford Conference, P. Losleben (ed.), Combridge: MIT Press, 295~312.

[7] 张立明, 人工神经网络的模型及其应用, 复旦大学出版社, 1993.

[8] 王东生、曹磊, 混沌、分形及其应用, 中国科学技术大学出版社, 1995.

[9] F. Pasemann & H. D. Doebner eds., Proceedings of the 9th Summer Workshop on Mathematical Physics, Arnold Sommerfeld Institute for Mathematical Physics Clausthal, 17~20, July, 1990.

[10] 阮炯, 具有非对称权阵的 Hopfield 连续神经网络的稳定性分析, 复旦学报, Vol.31, No.2(1992), 227~233.

[11] 阮炯, 具时滞的非线性神经网络连续系统的混沌, 复旦学报, Vol.33, No.3(1994).

[12] 赵维锐、阮炯, 具时滞的连续神经网络的稳定性分析, 复旦学报, Vol.33, No.2(1994).

[13] 阮炯, 具有延时的神经网络的稳定动力学行为分析, 复旦学报, Vol.34, No.2(1995).

[14] 顾凡及、李训经、阮炯, 动态神经元的网络模型 I, 生物物理学报, 2(1992).

[15] 顾凡及、李训经、阮炯, 动态神经元的网络模型 II, 生物物理学报, 3(1992).

[16] Ruan Jiong, A global stable analysis for CGNN and CNN with asymmetric weights, 日本名古屋国际神经网络大会论文集, 1993.10.

[17] 顾凡及、阮炯, 动态神经元网络模型的复杂性问题, 生物物理学报, Vol.33, No.2(1994).

[18] A. L. Hodgkin & A. F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, J. Physiol, 117(1952), 500~544.

[19] T. R. Chay, Chaos in a three–variable model of an excitable cell, Physica. D., 16(1985), 233~242.

[20] A. V. Holden, Models of the stochastic activity of neurones, Lecture notes in biomathematics, 12, Berlin: Springer–Verlag, 1976.

[21] J. Guckenheimer, J. Moser & S. Newhouse, Dynamical systems, V. Coates and S. Helgason, Eds. Birkhäuser, 1980.

[22] B. Hassard, Bifurcation of periodic solutions of the Hodgkin–Huxley model for the squid giant axon, J. Theoret. Biol., 71(1978), 401~420.

[23] J. Rinzel, On repetative activity in nerve, Federation Proc., 37(1978), 2793~2802.

[24] R. B. Stein, The frequency of nerve action potentials generated by applied currents, Proc. Roy. Soc. London, B167(1967), 64~86.

[25] N. H. Sabah & R. A. Spangler, Repetitive response of the Hodgkin–Huxley model for the squid giant axon, J. Theoret. Biol., 29(1970), 155~171.

[26] J. Rinzel & R. N. Miller, Numerical calculation of stable and unstable periodic solutions to the Hodgkin–Huxley equations, Math. Biosci., 49(1980), 27~59.

[27] E. N. Best, Null space in the Hodgkin–Huxley equations: A critical test, Biophys. J., 27(1979), 87~104.

[28] R. Guttman, S. Lewis & J. Rinzel, Control of repetitive firing in squid axon membrane as a model of a neurone oscillator, J. Physiol, 305(1980), 377~395.

[29] A. V. Holden & M. Yoda, Ionic channel density of an excitable membrane can act as a bifurcation parameter, Biol. Cybern, 42(1981), 29~38.

[30] A. V. Holden & M. Yoda, Bifurcation theory and autorhythmicity of the excitable membrane of nerve cells, in Proc. 2nd World COngr. on Mathematics on the Service of Man, A. Ballester, D. Cardus and E. Trillas, Eds. Las Palmas: Universidad Politecnica, 1982, 355~360.

[31] E. F. Mischenko & N. Kh. Rozov, Differential equations with small parameter and relaxation oscillations, Translated from Russian, London: Plennum, 1980.

[32] R. Lorente de No, A study of nerve Physiology, Stud. Rockerfeller Inst., 132(1947), 384~476.

[33] Abnormal Neuronal Discharges, N. Chalazonitis and M. Boisson, Eds. New York: Raven Press, 1978.

[34] G. A. Carpenter, Bursting phenomena in excitable membranes, SIAM. J. Appl. Maths., 36(1979), 334~372.

[35] C. K. Peterson, Techniques to trace curves of bifurcation points of perodic solutions, Path User's Guide, Centre for nonlinear studies, Leeds, 1987.

[36] A. V. Holden & M. A. Muhamad, The identification of deterministic chaos in the activity of single neurones, J. Electrophysiol. Techniques, 11, 135~147.

[37] N. H. Packard, J. P. Crutchfield, D. Farmer & R. S. Shaw, Geometry from a time-series, Phys. Rev. Letts., 45(1980), 712~716.

[38] D. Ruelle, Strange attractors, Math. Intelligencer, 3(1980), 126~137.

[39] D. Broomhead & G. King, Extracting qualitative dynamics from experimental data, Physica. D., 20(1986), 217.

[40] Jiong Ruan & Wei Lin, Chaos in a class of impusive differential equations, Com. Non. Sci & Num. Sim., 4:2(1999), 165.

[41] E. R. Caianiello & A. de Luca, Decision equation for binary systems: application to neural behavior, Kybernetik, 3(1966), 33～40.

[42] K. Gopalsamy & I. K. C. Leung, Convergence under dynamical thresholds with delays, IEEE Trans. Neural Networks, 8:2(1997), 341～348.

[43] K. Pakdaman & C. P. Malta, A note on convergence under dynamical thresholds with delays, IEEE Trans. Neural Networks, 9:1(1998), 231～233.

第四章　神经网络模型的应用

前面各章介绍了多种神经网络模型的背景及相应的算法. 本章介绍神经网络模型在生物医学工程、信号处理、模式识别等领域中的应用.

§4.1　生物医学工程

4.1.1　急性心肌梗塞的预后

近年来美国每年有 35~40 万人、德国有 8~10 万人发生心脏性猝死, 我国心脏性猝死的绝对数字也不小, 但目前尚无准确的统计数字. 猝死患者最常见的基础心脏病是冠心病, 该疾病的发生率正在逐年增长. 因此寻找一种无创的、准确的方法诊断心脏性猝死、预测急性心肌梗塞后的危险因子是人们的迫切要求, 也是医学界关注的热点之一.

心率变异 (heart rate variability, 简称 HRV) 是医学界普遍承认的无创性诊断急性心肌梗塞预后的重要指标之一.

心脏搏动是一种自主神经活动, 它反映了一个人心脏功能的健康状况. 对每个人来说, 心脏搏动的频率并不是绝对规则的, 两次心搏周期之间存在着微小的时间差异, 称为 **心率变异** (HRV). HRV 是一个检测自主神经活动的无创性指标, 对预测心血管疾病, 特别是急性心肌梗塞的预后, 具有重要的意义.

20 世纪 70 年代以来, 随着动态心电图 (holter) 记录技术的进步和数字信号处理方法的应用, HRV 的研究取得了很大的进展. 常用的分析方法是时域统计和频域功率谱方法, 这两种方法本质上都是线性方法. 1994 年, 欧洲心血管病学会与北美起搏和电生理学会组成了专题委员会, 讨论并制定了 HRV 的标准化时域和频

域指标, 以指导和推广线性指标在临床上的应用. 但新近的研究发现多种生理因素导致的心率总变化是非线性的变化, 不是各个单因素作用的简单叠加. 由于 HRV 受到血液动力、电生理、体液及自主神经调控之间的复杂的相互影响, 具有非线性的性质, 仅用线性方法对其进行研究, 可能会产生较大的偏差, 需要寻找新的途径来解决问题.

这里介绍一种用 **BP 网络** 对时间序列进行预测分析的方法.

首先介绍 BP 网络的基本算法.

设输入向量为 $x = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, 输出向量为 $y = (y_1, y_2, \cdots, y_m)^{\mathrm{T}}$, 连接第 i 个输入单元与第 j 个输出单元的权系数为 w_{ij} $(1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant n)$, 输入与输出的关系为

$$y_i = f\left(\sum_{j=1}^{n} w_{ij} x_j - \theta_i\right) \quad (1 \leqslant i \leqslant m), \qquad (4.1.1)$$

其中 θ_i $(1 \leqslant i \leqslant m)$ 为阈值, 而 $f(u)$ 为非线性函数, 通常取为 Sigmoid 函数

$$f(u) = \frac{1}{1 + \mathrm{e}^{-u}}. \qquad (4.1.2)$$

单层 BP 网络的示意图见图 4.1.1.

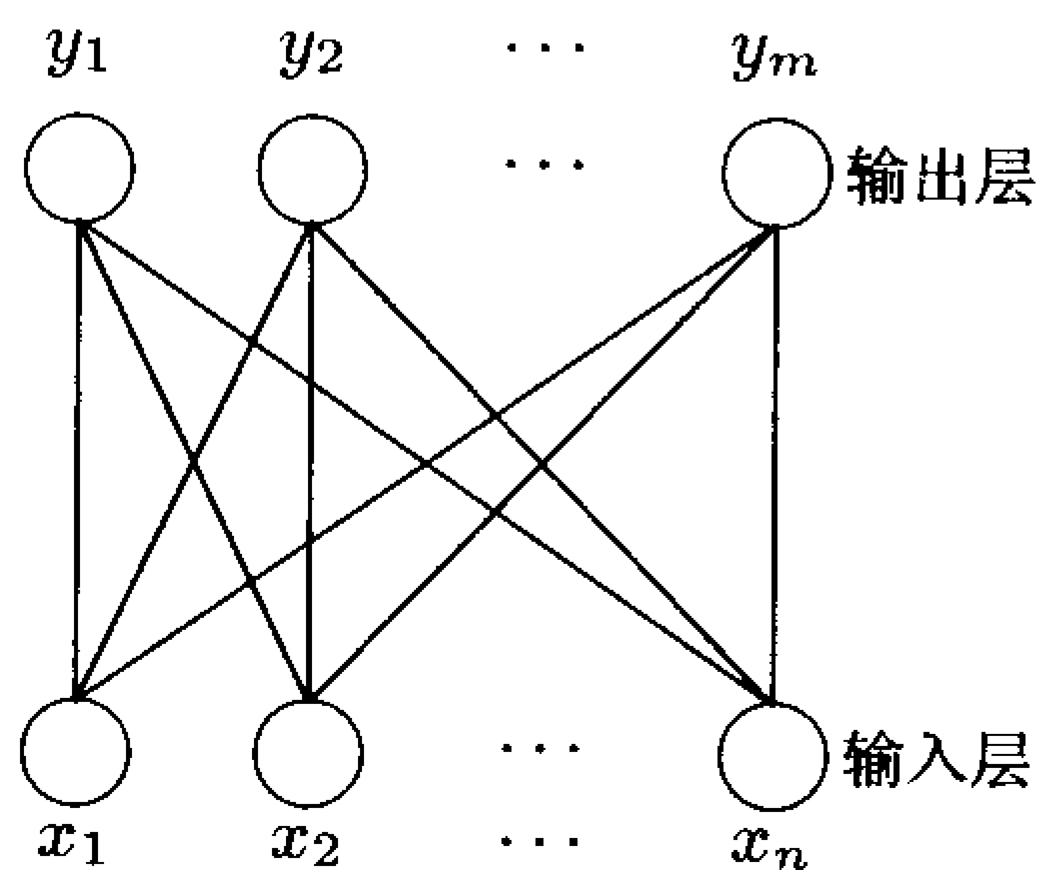


图 4.1.1 单层 BP 网络示意图

为方便起见, 记 $w_{i0} = \theta_i$ $(1 \leqslant i \leqslant m)$, $x_0 = -1$, 则 (4.1.1) 式可改写为

$$y_i = f\left(\sum_{j=0}^{n} w_{ij} x_j\right) \quad (1 \leqslant i \leqslant m). \qquad (4.1.3)$$

BP 算法是一种有教师的学习算法. 如果已知 BP 网络的要求响应 (即教师) 为 $t=(t_1,t_2,\cdots,t_m)^{\mathrm{T}}$, 则学习算法的目的是调节权系数 $w=(w_{ij})$, 使得网络的实际输出向量 y 与要求响应 t 尽可能接近.

设有 k 个学习样本, 输入样本为 $x^1,x^2,\cdots,x^k$, 相应的要求响应为 $t^1,t^2,\cdots,t^k$. 由 (4.1.3) 式, 实际输出 $y^1,y^2,\cdots,y^k$ 与要求响应 $t^1,t^2,\cdots,t^k$ 之间的平方误差为

$$
\begin{aligned}
\varepsilon &= \frac{1}{2}\sum_{l=1}^{k}\sum_{i=1}^{m}(y_i^l-t_i^l)^2 \\
&= \frac{1}{2}\sum_{l=1}^{k}\sum_{i=1}^{m}\left[f\left(\sum_{j=0}^{n}w_{ij}x_j^l\right)-t_i^l\right]^2,
\end{aligned}
\tag{4.1.4}
$$

这里取系数 $\frac{1}{2}$ 仅仅是为了计算上的方便.

我们的目标是求权系数 w_{ij}, 使误差 ε 达到最小值. 这是一个无约束的极值问题, 可以采用梯度法来求解. 此时权系数 w 的修正值沿 ε 关于 w 的梯度的反方向, 也就是使 ε 减小最快的方向, 即

$$
\begin{aligned}
\Delta w_{ij} &= -\eta\frac{\partial\varepsilon}{\partial w_{ij}} \\
&= -\eta\sum_{l=1}^{k}(y_i^l-t_i^l)f'(u_i^l)x_j^l,
\end{aligned}
\tag{4.1.5}
$$

其中 η 为学习率参数, 而

$$
u_i^l=\sum_{j=0}^{n}w_{ij}x_j^l \quad (1\leqslant i\leqslant m, 1\leqslant l\leqslant k). \tag{4.1.6}
$$

用梯度法可以使权系数 w 沿着误差 ε 减小最快的方向变化, 直到 $\Delta\varepsilon=0$. 这种学习方法可以使权系数 w_{ij} 稳定到一个解, 它一般是误差 ε 的一个局部极小解, 但不能保证是全局最小解.

类似地对于多层 BP 网络, 设输入向量为 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, 输出向量为 $y=(y_1,y_2,\cdots,y_m)^{\mathrm{T}}$, 中间加入 k 层隐层, 第 l 层隐层有 n_l 个神经元, 相应的神经元向量为 $x^{(l)}=(x_1^{(l)},x_2^{(l)},\cdots,x_{n_l}^{(l)})^{\mathrm{T}}$. 为方便起见, 记 $x^{(0)}=x$, $x^{(k+1)}=y$, 则各层神经元的输出之间满足

$$x_i^{(l+1)}=f\left(\sum_{j=1}^{n_l}w_{ij}^{(l)}x_j^{(l)}-\theta_i^{(l)}\right)\quad(0\leqslant l\leqslant k),\tag{4.1.7}$$

其中 $w_{ij}^{(l)}$ 为权系数, $\theta_i^{(l)}$ 为阈值. 多层 BP 网络示意图如图 4.1.2 所示.

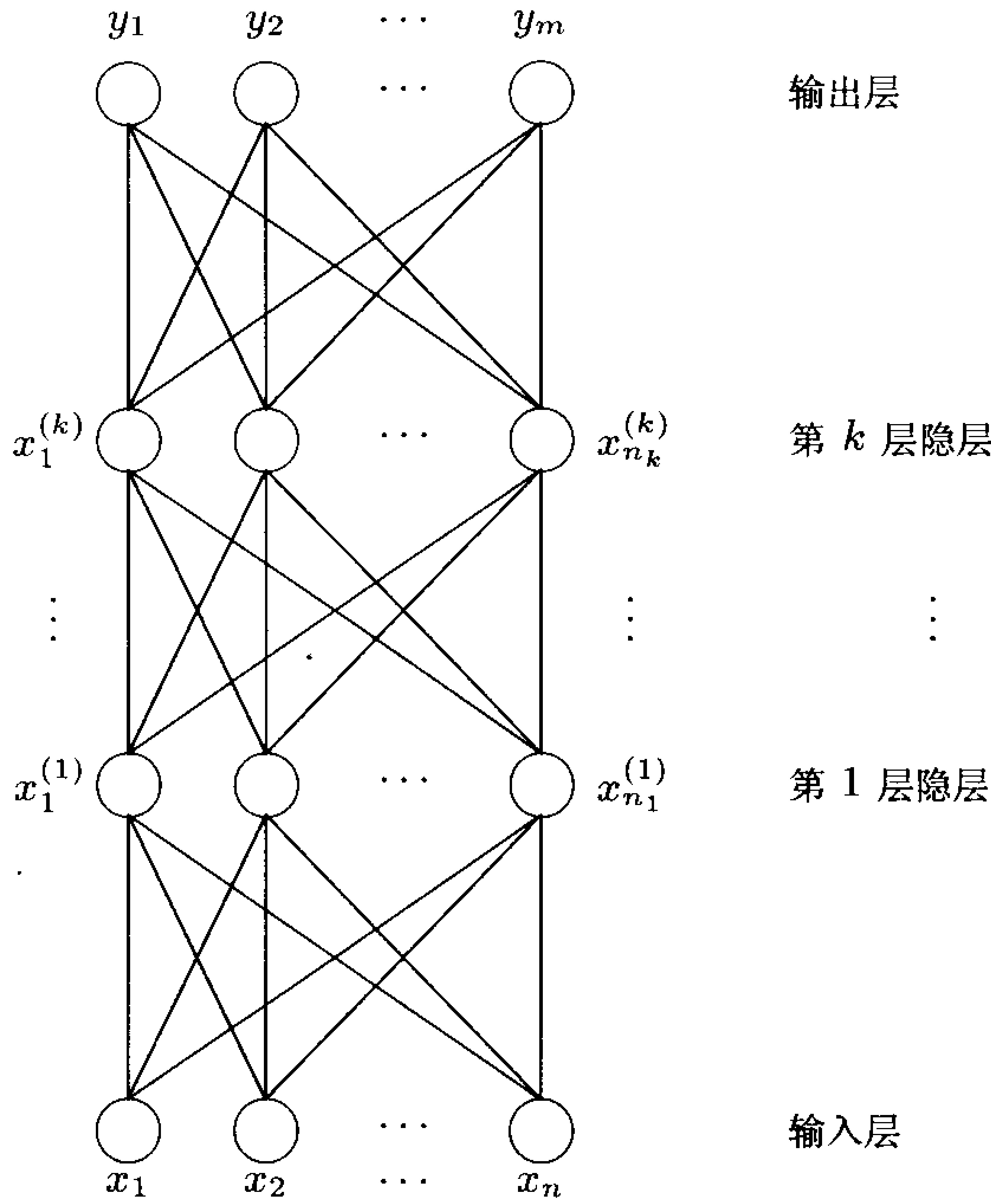


图 4.1.2　多层 BP 网络示意图

同单层 BP 网络一样, 令 $w_{i0}^{(l)}=\theta_i^{(l)}$, $x_0^{(l)}=-1$, 则 (4.1.7) 式

可改写为

$$x_i^{(l+1)} = f\left(\sum_{j=0}^{n_l} w_{ij}^{(l)} x_j^{(l)}\right) \quad (0 \leqslant l \leqslant k). \tag{4.1.8}$$

若已知网络的要求响应为 $t=(t_1,t_2,\cdots,t_m)^{\mathrm{T}}$, 则我们的目的是调节权系数 $w_{ij}^{(l)}$, 使得网络的实际输出 y 与要求响应 t 之间的平方误差达到最小. 采用梯度法, 类似于 (4.1.5) 式的推导, 可以得到 $w_{ij}^{(l)}$ 的修正值 $\Delta w_{ij}^{(l)}$.

这样, 整个网络的学习过程可以描述如下:

(1) 随机给出各层的初始权系数 $w_{ij}^{(l)}$ 和阈值 $\theta_i^{(l)}$;

(2) 对于每个学习样本由 (4.1.7) 或 (4.1.8) 式计算网络的实际输出 y, 并计算相应的误差 ε;

(3) 由梯度法得到权系数和阈值的修正值;

(4) 根据新的权系数和阈值重新计算网络的实际输出和误差 ε. 当误差 ε 充分小时, 学习结束; 否则重复步骤 (3), 重新修改权系数和阈值, 直到误差 ε 很小为止.

下面我们用 BP 网络对心率的 R–R 间期进行分析. 设 $x=(x_1,\cdots,x_n,x_{n+1},\cdots,x_{2n})^{\mathrm{T}}$ 是某个人在一定时间内的心率周期, 可以视为一个时间序列. 将该序列分为两组: $x_1,\cdots,x_n$ 为学习组, $x_{n+1},\cdots,x_{2n}$ 为检验组. 利用学习组的时间序列通过 BP 算法得到一组最佳权系数值, 然后将这组权系数用于检验组的序列, 得到检验组序列的预测值. 将预测值与原检验组序列的真实值加以比较, 得到它们的相关系数, 作为该时间序列的预测度, 并以此分析其动力学性质.

可以证明: 周期的、有规律的时间序列单步和多步都是可预测的, 其预测度均较高. 随机的、无规律的时间序列单步和多步都是不可预测的, 其预测度均较低. 而具有混沌性质的时间序列则是单步可预测, 随着预测步长的增加, 逐渐向不可预测转化, 其预测度开始较高, 逐渐递减.

我们用这种基于 BP 网络的预测度方法对上海市第一人民医院提供的 51 例急性心肌梗塞 (AMI) 、 ST 段下移等患者和 44 例

正常对照组的心率变异进行分析对比，结果如表 4.1.1 所示

表 4.1.1

	正常对照组	AMI 等患者
可预测	2	6
不可预测	7	31
混沌	35	14
合计	44	51

从上表可以看出：对健康者来说，他们的心率变异多呈混沌状，用预测度方法分析的结果为单步的预测度较高，而多步的预测度逐渐降低；而对 AMI 等心肌疾病患者来说，他们的心率变异多为可预测的或不可预测的.

上海市第一人民医院的跟踪结果显示，AMI 患者中有 4 人已死亡，用预测度指标进行分析，他们的心率变异均为不可预测的；而用线性指标进行分析，只有 1 人显示出有较严重的疾病. 由此可见，预测度指标较线性指标有更好的诊断能力，是一种更好的临床诊断指标.

4.1.2 胎儿心率的检测

计划生育是我国的基本国策之一，现在我国大多数夫妇都只有一个孩子，孩子的优生优育是所有家庭都关注的问题. 孩子生长的好坏与其在胎儿时在母体内的发育有很大的关系，母亲的孕期检查尤为重要. 孕期检查中有一项是检查胎儿的心率情况. 由于母亲的心率很强，在测量胎儿的心率时，不可避免地会受到母体心率的干扰. 如何去除母体心率，得到胎儿的真实心率信息是一个非常困难的问题，通过滤波器是很难做到的. 考虑到胎儿心率与母体心率是相互独立的，我们希望通过自适应线性神经网络，将母体心率作为噪声源，而把混有噪声的胎儿心率信号作为要求的响应，通过对消得到清晰的胎儿心率.

自适应线性神经网络 (Adaline 网络) 是 Standford 大学的

Windrow 教授于 1961 年首先提出的. 设输入向量为 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, 权系数向量为 $w=(w_1,w_2,\cdots,w_n)^{\mathrm{T}}$, 输出分为两个部分: 模拟输出部分 y 和数字输出部分 z. 输入与输出之间的关系为

$$y=\sum_{i=1}^{n}w_ix_i-\theta, \tag{4.1.9}$$

$$z=\operatorname{sgn}(y), \tag{4.1.10}$$

其中 θ 为阈值, $\operatorname{sgn}(y)$ 为符号函数

$$\operatorname{sgn}(y)=\begin{cases}1, & y\geqslant 0,\\ -1, & y<0.\end{cases}$$

自适应线性神经网络的示意图参见图 4.1.3.

令 $w_0=\theta$, $x_0=-1$, 则 (4.1.9) 式可改写为

$$y=\sum_{i=0}^{n}w_ix_i. \tag{4.1.11}$$

如果已知 Adaline 网络的要求响应为 t, 则我们的目的是调节权系数向量 w, 使得网络的模拟输出向量 y 与要求响应 t 之间的误差达到最小. 显然 Adaline 网络也是一种有教师的学习算法. 通常采用下面的 LMS 算法, 它保证 Adaline 网络的学习过程是收敛的.

设当前迭代为第 k 次迭代, $x^k=(x_1^k,x_2^k,\cdots,x_n^k)$ 为当前输入样本, ε^k 为当前模拟输出 y^k 与要求响应 t 之间的误差, $w^k=(w_1^k,\cdots,w_n^k)$ 为当前的权系数向量, w^{k+1} 为下一次迭代的权系数向量, 则

$$w^{k+1}=w^k+\frac{\alpha}{\|x^k\|^2}\varepsilon^kx^k, \tag{4.1.12}$$

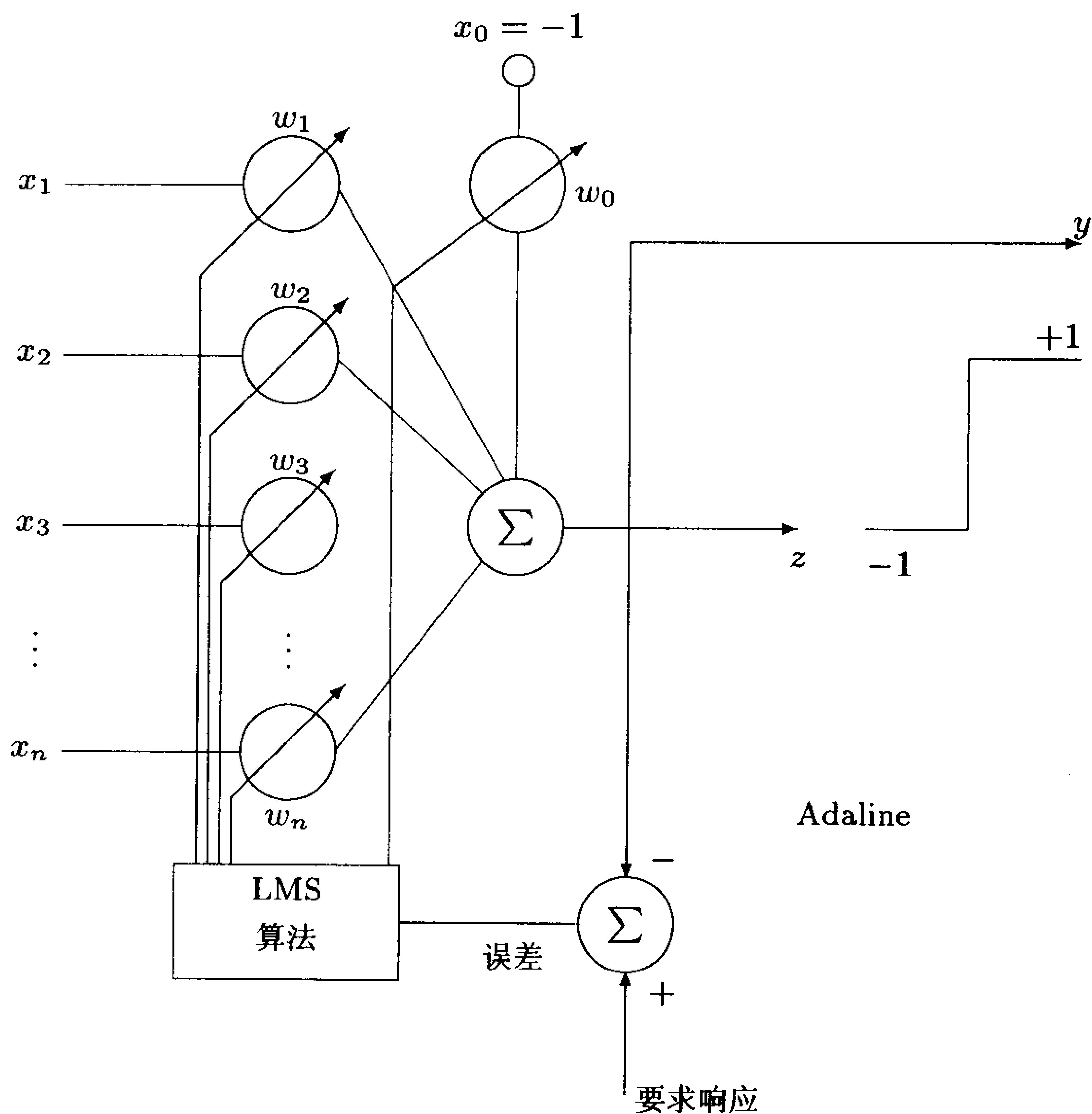


图 4.1.3　自适应线性神经网络示意图

其中 α 是控制因子. 当前模拟输出与要求响应之间的误差为

$$\varepsilon^k = t - y^k = t - \sum_{i=1}^{n} w_i^k x_i^k = t - (w^k)^{\mathrm{T}} x^k. \tag{4.1.13}$$

当网络的输入不变，而权系数改变时，相应的误差改变为

$$\Delta\varepsilon^k = \Delta[t - (w^k)^{\mathrm{T}} x^k] = -(\Delta w^k)^{\mathrm{T}} x^k, \tag{4.1.14}$$

$$\Delta w^k = w^{k+1} - w^k = \frac{\alpha}{\|x^k\|^2} \varepsilon^k x^k. \tag{4.1.15}$$

将 (4.1.15) 式代入 (4.1.14) 式，得到

$$\Delta\varepsilon^k = -\frac{\alpha}{\|x^k\|^2} \varepsilon^k (x^k)^{\mathrm{T}} x^k = -\alpha\varepsilon^k. \tag{4.1.16}$$

这说明，当网络的输入不变时，误差的变化为当前误差的 α 倍. 又由

$$\varepsilon^{k+1} = (1-\alpha)\varepsilon^k,$$

下一次迭代的误差为当前误差的 $(1-\alpha)$ 倍. 因此适当地选取 α 因子可以控制收敛的速度和稳定性. 显然当 $|1-\alpha| < 1$, 即 $0 < \alpha < 2$ 时， LMS 算法是收敛的，通常取 $0.1 < \alpha < 1.0$.

用 Adaline 网络进行噪声对消，其示意图如图 4.1.4, 其中 AD 表示 Adaline 网络. 设有一个输入信号 s, 加上一个与 s 不相关的噪声 n_0 作为 Adaline 网络的要求响应，而将与噪声 n_0 相关的信号 n_1 输入 Adaline 网络组成的自适应滤波器，其输出为 y, y 与 n_1 相关. 于是网络的输出与要求响应之间的误差为

$$\varepsilon = s + n_0 - y. \tag{4.1.17}$$

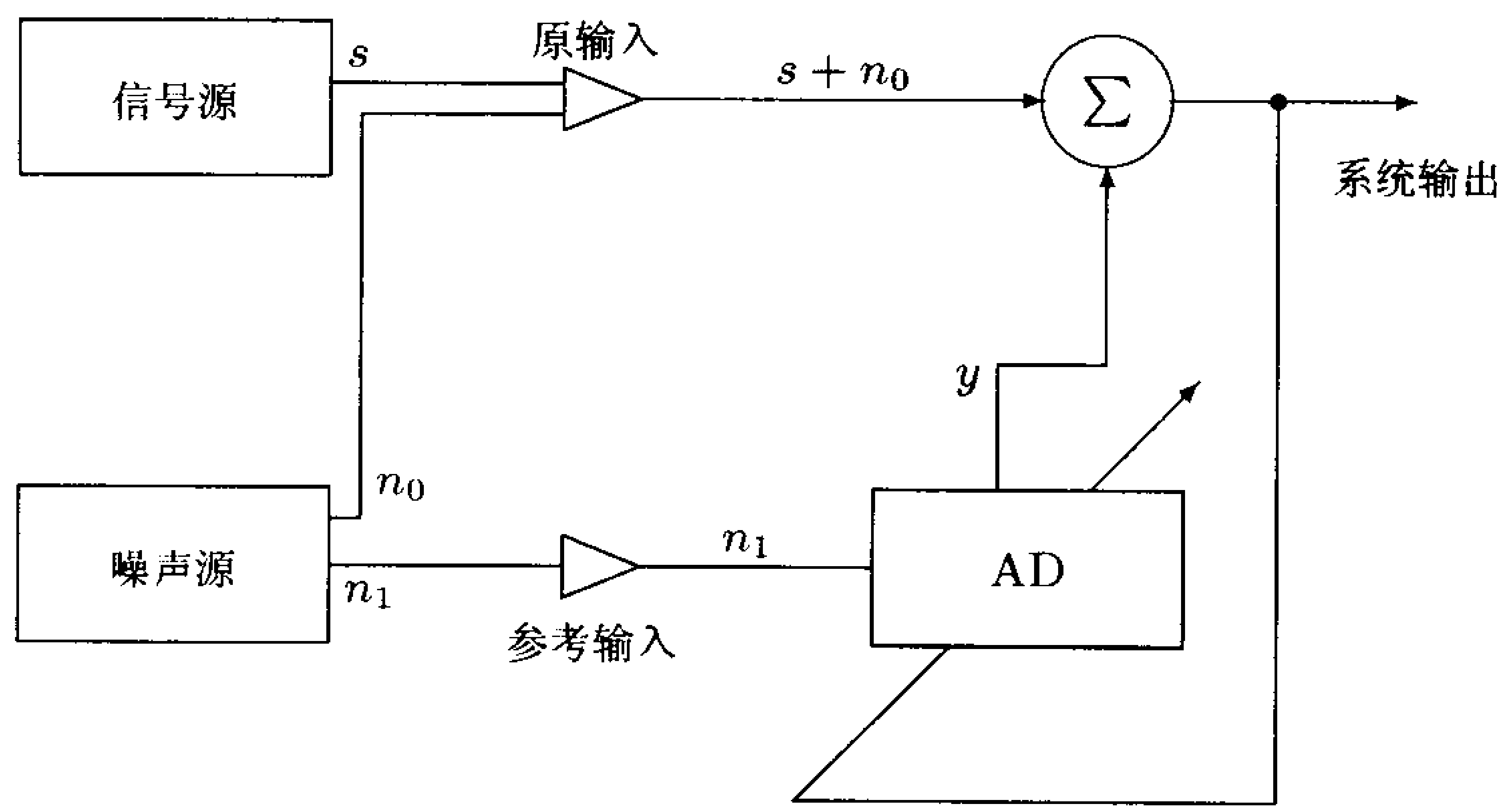


图 4.1.4　噪声对消示意图

如果 s 是均值为 0 的平稳随机信号，而 n_0 是一个独立于 s 的随机噪声，那么可以得到

$$\varepsilon^2 = s^2 + (n_0 - y)^2 + 2s(n_0 - y), \tag{4.1.18}$$

$$\begin{aligned} E(\varepsilon^2) &= E(s^2) + E(n_0 - y)^2 + 2E(s(n_0 - y)) \\ &= E(s^2) + E(n_0 - y)^2, \end{aligned} \tag{4.1.19}$$

其中 $E(X)$ 表示随机变量 X 的均值.

通过调节 Adaline 网络的权系数, 得到

$$E_{\min}(\varepsilon^2) = E_{\min}(s^2) + E_{\min}(n_0 - y)^2. \tag{4.1.20}$$

当 $E_{\min}(n_0 - y)^2 \to 0$ 时, y 趋近于 n_0. 此时系统的输出 ε 即为 s, 其噪声 n_0 被抵消了.

采用这种系统来检测胎儿的心率, 效果非常好.

§4.2 信 号 处 理

4.2.1 数据压缩

现代社会是信息社会, 信息在社会经济生活中占有极其重要的地位. 谁在经营活动中掌握了信息, 谁就把握住了机会; 谁在战争中获得了信息, 谁就取得了胜机. 通常信息表现为数据, 即使是非数据信息, 考虑到信息之间的交流, 一般也将其量化. 由于信息量通常非常庞大, 信息在交流时所需要传输的数据量也就非常可观, 同时存放这些信息也需要大量的存贮单元. 因此如何有效地压缩数据, 尽可能减少数据的存贮量和传输的费用就成为许多科学家的研究课题.

数据压缩又称为数据的特征提取. 将真实数据用较少的单元来存贮和传递, 使用时再将压缩数据还原成原始数据, 这样就大大节省了存放和传送的开销. 这里采用多层 BP 网络来完成这项工作.

BP 网络已在 §4.1.1 中详细介绍过. 这里用于数据压缩的 BP 网络由 3 层神经元组成, 输入层与输出层的神经元均为 n 个, 中间层的神经元为 $n_1 = \log_2 n$ 个 (参见图 4.2.1).

当输入各个学习样本后, 通过 BP 网络的学习使输入和输出样本完全相同. 由于中间层的神经元数远远小于输入和输出层的神经元数, 中间层的神经元便成为输入样本的一种压缩. 将中间层的神经元的输出存放在内存里, 使用时只需乘上相应的权系数就

可以恢复原来的数据. 在信息存贮时, 可存贮这些压缩数据, 减少存贮单元的费用. 在信息传输时, 可以发出压缩数据, 在接收时乘上相应的权, 即可恢复原来的信息.

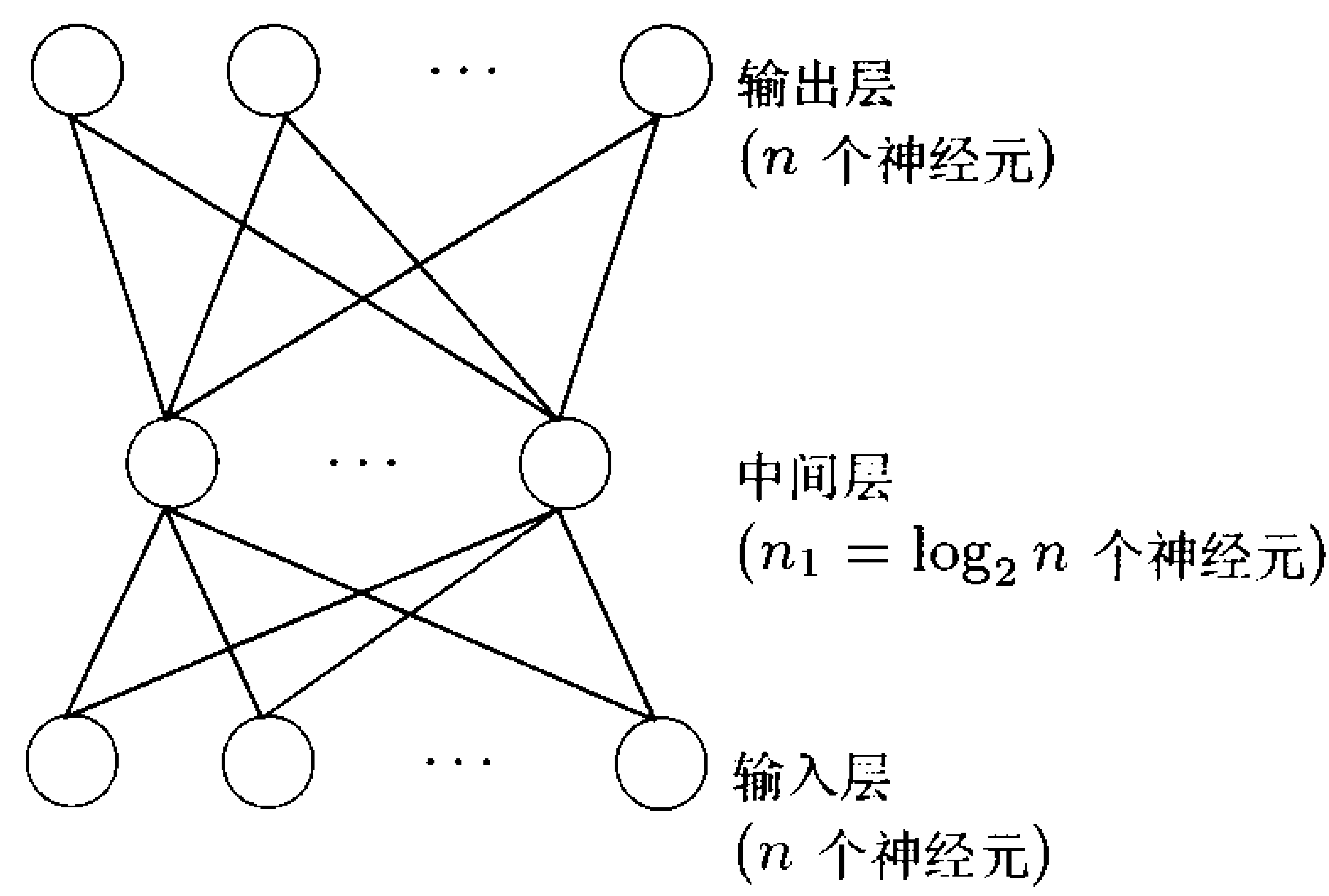


图 4.2.1 用于数据压缩的 3 层 BP 网络示意图

表 4.2.1 是一个数据压缩的例子. 输入样本由一串 8 位二进制数构成, 其中只有一位为 1, 其余均为 0. 由上述 3 层 BP 网络, 中间层采用 $\log_2 8 = 3$ 个神经元. 在计算机内部存储时, 只要存入中间神经元的输出, 使用时就可以很快恢复原来的代码.

表 4.2.1 数据压缩的 3 层 BP 网络的各层输出值

输 入	中 间			输 出
1 0 0 0 0 0 0 0	0.5	0	0	1 0 0 0 0 0 0 0
0 1 0 0 0 0 0 0	0	1	0	0 1 0 0 0 0 0 0
0 0 1 0 0 0 0 0	1	1	0	0 0 1 0 0 0 0 0
0 0 0 1 0 0 0 0	1	1	0	0 0 0 1 0 0 0 0
0 0 0 0 1 0 0 0	0	1	1	0 0 0 0 1 0 0 0
0 0 0 0 0 1 0 0	0.5	0	1	0 0 0 0 0 1 0 0
0 0 0 0 0 0 1 0	1	0	0.5	0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 1	0	0	0.5	0 0 0 0 0 0 0 1

4.2.2 保密通信

当前建立信息高速公路是各国计算机产业的一大热点，国际互连网的普及，全球计算机网络化，大大缩短了世界各地之间的距离. 在信息互通、资源共享的同时，一些重要信息的保密安全问题也提到了议事日程上. 我们介绍一种用动态神经网络对数据进行加密和解密的方法.

动态神经网络在 §3.9 中已经有较详细的介绍. 现给出它的离散化的计算公式

$$x_{ij}(t+\Delta t)=\begin{cases} b(t_{ij})x_{ij}(t), & \text{若 } z_j(t-\tau_{ij})=0, \\ b(t_{ij})x_{ij}(t)+A_{ij}, & \text{若 } z_j(t-\tau_{ij})=0, \end{cases} \tag{4.2.1}$$

$$x_{i0}(t+\Delta t)=A_{i0}I_i(t)\Delta t, \tag{4.2.2}$$

$$y_i(t+\Delta t)=\begin{cases} b(t_i)y_i(t)+\sum w_{ij}x_{ij}(t)\Delta t, & \text{若 } z_i(t)=0, \\ b(t_i)y_i(t)+\sum w_{ij}x_{ij}(t)\Delta t-\theta_i, & \text{若 } z_i(t)=1, \end{cases} \tag{4.2.3}$$

$$z_i(t)=\begin{cases} 0, & \text{若 } y_i(t)<\theta_i, \\ 1, & \text{若 } y_i(t)\geqslant\theta_i, \end{cases} \tag{4.2.4}$$

$$b(t)=\mathrm{e}^{-\frac{\Delta t}{t}}, \tag{4.2.5}$$

其中 t_{ij}, t_{i0}, t_i 分别为神经元各处膜的时间常数， y_i 为始段分级电位， x_{ij} 为突触后电位， x_{i0} 为感受器电位， z_i 为轴突上的动作电位序列， τ_{ij} 为突触延搁， θ_i 为阈值. 我们引进一个函数表示阈值 $\theta_i(t)$

$$\theta_i(t)=\begin{cases} \infty, & \text{若 } t-t_i<1, \\ \theta_0+a_\theta\mathrm{e}^{-\frac{t-t_i}{t\theta_1}}\cos\left(\mathrm{e}^{-\frac{t-t_i}{t\theta_2}}\dfrac{t-t_3}{t\theta_3}\right), & \text{若 } t-t_i\geqslant 1, \end{cases} \tag{4.2.6}$$

其中 t_i 是第 i 个神经元最近一次发动作电位的时间.

这个模型可用来产生类似脑电波这样的复杂波形. 在这个模型中没有引进任何噪声, 而通过数值模拟, 此模型的状态变量突触后电位的值对初值非常敏感. 例如在图 4.2.2 中, 把计算上一条曲线 (a) 的一个神经元突触后电位的初值仅仅改变了 0.01, 重新算得的 y 值由下一条曲线 (b) 所示, 它们的差由曲线 (c) 所示.

由于混沌的一条基本性质是对初值的极端敏感性, 因此这个模型具有高度复杂的混沌特性. 当然也可以通过计算该系统的 Lyapunov 指数或复杂度这类非线性混沌动力学指标来判断该系统的混沌特性. 我们将其用于信息的加密.

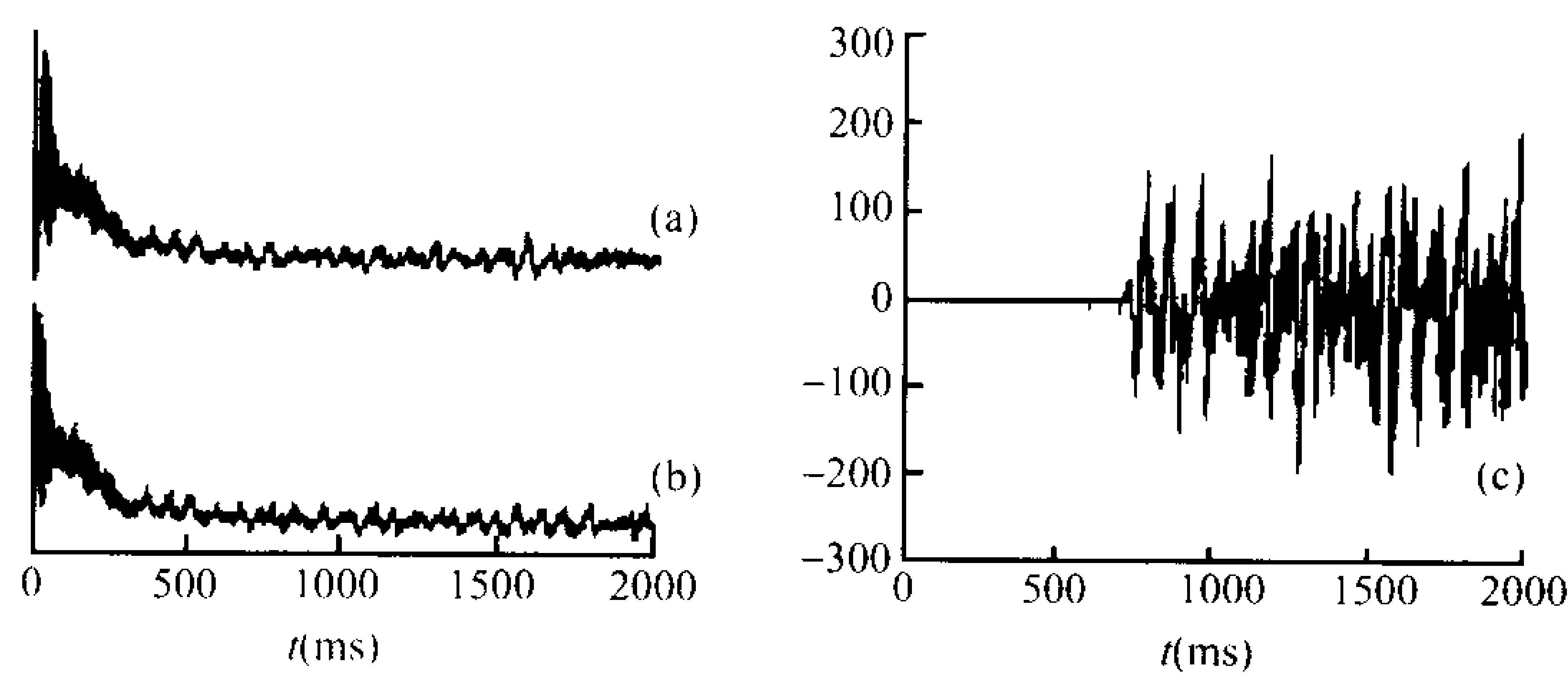


图 4.2.2　动态神经网络对初值的极端敏感性

首先将待加密的信息看作一串二进制数字序列 A, 然后在适当的参数配置下由混沌神经网络产生一段时间序列, 并将它也转换成一串二进制数字序列 B, 称为 **信号载体**. 将原始序列 A 与信号载体 B 作按位的异或 (XOR) 运算, 得到一串二进制数字序列 C, 这就是经过加密的信息. 将这串信息通过计算机网络发送给对方, 在恢复信息时, 只要用相同的参数重新产生信号载体 B 序列, 再对 B 序列与 C 序列作按位的异或 (XOR) 运算, 就可以重新得到原始序列 A.

由于此模型对初值极端敏感, 因此初值的微小扰动会引起网络输出的很大变化, 使得解密失败, 从而达到保密的作用.

图 4.2.3 是用此模型进行加密的简单例子. 图 (a) 是原始文字信息, 图 (b) 是加密以后的文字信息, 图 (c) 是用相同的初值得到

的序列进行解密的结果，与原始信息完全相同，图 (d) 是用几乎相同，仅有一位有 0.001 的误差的初始值得到的序列进行解密的结果，解密失败.

(a) Hello, World!

(b) ┤╜ÆΓ▌╓╖█≠V²↕▀α⌠•

(c) Hello, World!

(d) ┌k•b]!↕ztvw'{↕↕]

图 4.2.3　用动态神经网络加密与解密的例子

§4.3　优 化 问 题

4.3.1　旅行推销商问题 —— 离散 Hopfield 网络

旅行推销商问题 (TSP 问题，travelling salesman problem) 是一个十分困难的优化问题. 问题的提法是：一个推销员去 n 个城市推销他的产品，从他所在的城市出发访问 $n-1$ 个城市，最后又回到原来的城市. 要求所有城市都要走到 (称为 **遍历**)，且只经过一次，并且使得推销员所经过的路程最短. 这是一个 NP 完全问题，当城市数目 n 比较小时，采用动态规划尚能求解，而当城市数 n 增加时，计算量就会大大增加，即使使用计算机也难以胜任这一工作.

这个问题可以采用 Hopfield 网络加以解决.

首先建立 TSP 问题的优化模型. 设推销员访问的城市为 $A, B, C, D, \cdots$ 城市与城市之间的距离分别为 d_{AB}, d_{AC}, d_{AD}, $\cdots$, $d_{BC}, d_{BD}, \cdots, d_{CD}, \cdots$. 记 $M=(m_{xi})$ 为关联矩阵，其中第一个下标表示城市 $A, B, C, D, \cdots$, 用 $x, y, \cdots$ 表示，第二个下标表示推销员访问城市的顺序，用 $i, j, \cdots$ 表示. 如果推销员第 i 个访问的城市是 x, 那么 $m_{xi}=1$, 否则 $m_{xi}=0$. 这样，关联矩阵的元素 $m_{xi}=0$ 或 1, 而推销员走过的总距离为

$$d=\frac{1}{2}\sum_{i}\sum_{\substack{x \\ x\neq y}}\sum_{y} d_{xy}m_{xi}(m_{y,i+1}+m_{y,i-1}), \tag{4.3.1}$$

其中 $m_{y,0} = m_{y,n}$, $m_{y,n+1} = m_{y,1}$.

例如有 6 个城市分别记为 A, B, C, D, E, F, 如果推销员从 E 城出发，访问城市的顺序依次为

$$E \to C \to A \to F \to D \to B \to E,$$

那么相应的关联矩阵为

$$M = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

而推销员所走的距离为

$$d = d_{EC} + d_{CA} + d_{AF} + d_{FD} + d_{DB} + d_{BE}.$$

为了保证每个城市只访问一次，关联矩阵上每一行只能有一个元素为 1, 其余均为 0, 即关联矩阵应满足

$$\sum_x \sum_{\substack{i \\ i \neq j}} \sum_j m_{xi} m_{xj} = 0. \tag{4.3.2}$$

而每次推销员只能访问一个城市，关联矩阵的每一列也只能有一个元素为 1, 其余为 0, 因此关联矩阵还应满足

$$\sum_i \sum_{\substack{x \\ x \neq y}} \sum_y m_{xi} m_{yi} = 0. \tag{4.3.3}$$

此外，每个城市均应被访问，所以总的 m_{xi} 为 1 的总数应为 n, 即

$$\sum_x \sum_i m_{xi} = n. \tag{4.3.4}$$

于是 TSP 问题归结为求关联矩阵 $M=(m_{xi})$, 使得

$$\min d=\frac{1}{2}\sum_{i}\sum_{\substack{x\\x\neq y}}\sum_{y}d_{xy}m_{xi}(m_{y,i+1}+m_{y,i-1}),\tag{4.3.5}$$

满足约束条件

$$\sum_{x}\sum_{\substack{i\\i\neq j}}\sum_{j}m_{xi}m_{xj}=0,\tag{4.3.6}$$

$$\sum_{i}\sum_{\substack{x\\x\neq y}}\sum_{y}m_{xi}m_{yi}=0,\tag{4.3.7}$$

$$\sum_{x}\sum_{i}m_{xi}=n,\tag{4.3.8}$$

$$m_{xi}=0\quad 或\quad 1\quad (1\leqslant x,i\leqslant n).\tag{4.3.9}$$

这是一个约束极值问题, 用 Lagrange 乘子法可化为无约束极值问题

$$E=\frac{P}{2}\sum_{i}\sum_{\substack{x\\x\neq y}}\sum_{y}d_{xy}m_{xi}(m_{y,i+1}+m_{y,i-1})+\frac{Q}{2}\sum_{i}\sum_{\substack{x\\x\neq y}}\sum_{y}m_{xi}m_{yi}$$
$$+\frac{S}{2}\sum_{x}\sum_{\substack{i\\i\neq j}}\sum_{j}m_{xi}m_{xj}+\frac{T}{2}\left(\sum_{x}\sum_{i}m_{xi}-n\right)^{2},\tag{4.3.10}$$

其中系数 P,Q,S,T 和距离 d_{xy} 都是大于零的常数.

下面我们讨论如何利用 Hopfield 网络解决这个优化问题.

离散 Hopfield 神经网络是一个离散时间系统. 设 $x(t)=(x_1(t),\cdots,x_n(t))^{\mathrm{T}}$, 其中 $x_i(t)$ $(1\leqslant i\leqslant n)$ 表示第 i 个神经元在 t 时刻的状态, 第 i 个神经元的下一时刻的状态定义为

$$x_i(t+1)=\operatorname{sgn}\left(\sum_{j=1}^{n}w_{ij}x_j(t)-\theta_i\right)\quad(1\leqslant i\leqslant n)\tag{4.3.11}$$

或

$$x_i(t+1)=\frac{1}{2}+\frac{1}{2}\mathrm{sgn}\left(\sum_{j=1}^{n}w_{ij}x_j(t)-\theta_i\right)\quad(1\leqslant i\leqslant n),\quad(4.3.12)$$

其中 θ_i 为第 i 个神经元的阈值，而 $\mathrm{sgn}(u)$ 为

$$\mathrm{sgn}(u)=\begin{cases}1, & u\geqslant 0,\\ -1, & u<0.\end{cases}\qquad(4.3.13)$$

系统 (4.3.11) 或 (4.3.12) 就是离散 Hopfield 神经网络. 易知每个状态 $x_i(t)$ 均为二值函数，取值为 $\{-1,1\}$ 或 $\{0,1\}$.

离散 Hopfield 神经网络有两种工作方式：异步工作方式和同步工作方式.

异步工作方式是指在任一时刻 t, 只有一个状态发生变化，而其余状态保持不变，即 (对系统 (4.3.11) 而言)

$$x_i(t+1)=\mathrm{sgn}\left(\sum_{j=1}^{n}w_{ij}x_j(t)-\theta_i\right),\qquad(4.3.14)$$

$$x_k(t+1)=x_k(t)\quad(k\neq i).\qquad(4.3.15)$$

同步工作方式是指在任一时刻 t, 部分状态发生改变，其中最重要的一种特殊情形是完全同步工作方式，即所有状态同时发生改变,

$$x_i(t+1)=\mathrm{sgn}\left(\sum_{j=1}^{n}w_{ij}x_j(t)-\theta_i\right)\quad(1\leqslant i\leqslant n).\qquad(4.3.16)$$

对离散 Hopfield 神经网络若能构造出能量函数 $E(t)$, 使之收敛，则系统 (4.3.11) 将收敛于一个稳定的状态

$$x_i=\mathrm{sgn}\left(\sum_{j=1}^{n}w_{ij}x_j-\theta_i\right)\quad(1\leqslant i\leqslant n).\qquad(4.3.17)$$

对 TSP 问题的优化模型 (4.3.5)~(4.3.9), 由于关联矩阵 M 的每个元素 m_{xi} 都是二值函数, 因此可将它们视为离散 Hopfield 神经网络中的状态变量. 将由 (4.3.10) 式定义的函数 E 作为相应的能量函数, 显然 E 满足 $E>0$. 取 m_{xi} 满足状态方程

$$\frac{\mathrm{d}m_{xi}}{\mathrm{d}t}=-\frac{\partial E}{\partial m_{xi}}, \tag{4.3.18}$$

则

$$\frac{\mathrm{d}E}{\mathrm{d}t}=\sum_{x,i}\frac{\partial E}{\partial m_{xi}}\frac{\mathrm{d}m_{xi}}{\mathrm{d}t}=-\sum_{x,i}\left(\frac{\partial E}{\partial m_{xi}}\right)^2\leqslant 0. \tag{4.3.19}$$

因此能量函数 E 是一个单调下降有下界的函数, 它是收敛的. 由此系统将收敛到使 E 达到最小的一些稳定状态上, 它们即为所要求的较优路径的解.

4.3.2 TSP 问题 —— Kohonen 网络

前面利用 Hopfield 网络来解 TSP 问题, 下面我们利用自组织学习方法来求解 TSP 问题.

自组织映照模型 是由 T. Kohonen 提出的, 又称为 **Kohonen 网络**, 它与前面提到的 BP 网络、 Hopfield 网络的一个重要区别在于它是无教师的学习网络. Kohonen 网络由输入层和输出层两层神经网络组成. 输入层中的每一个神经元, 通过权与输出层的每一个神经元相连, 输出层中的神经元通常以二维形式排列 (参见图 4.3.1).

设输入向量为 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, 输出向量为 $y=(y_1, y_2,\cdots,y_m)^{\mathrm{T}}$, 第 i 个输入神经元 x_i 与第 j 个输出神经元 y_j 之间的权为 w_{ij}, 输入与输出满足状态方程

$$\frac{\mathrm{d}y_j}{\mathrm{d}t}=\sum_{i=1}^{n}w_{ij}x_i-f(y_j)\quad(1\leqslant j\leqslant m), \tag{4.3.20}$$

其中 $f(y_j)$ 是一个单调增加的非线性函数, 前面的负号表示 $f(y_j)$ 的作用是减缓 y_j 的变化速度.

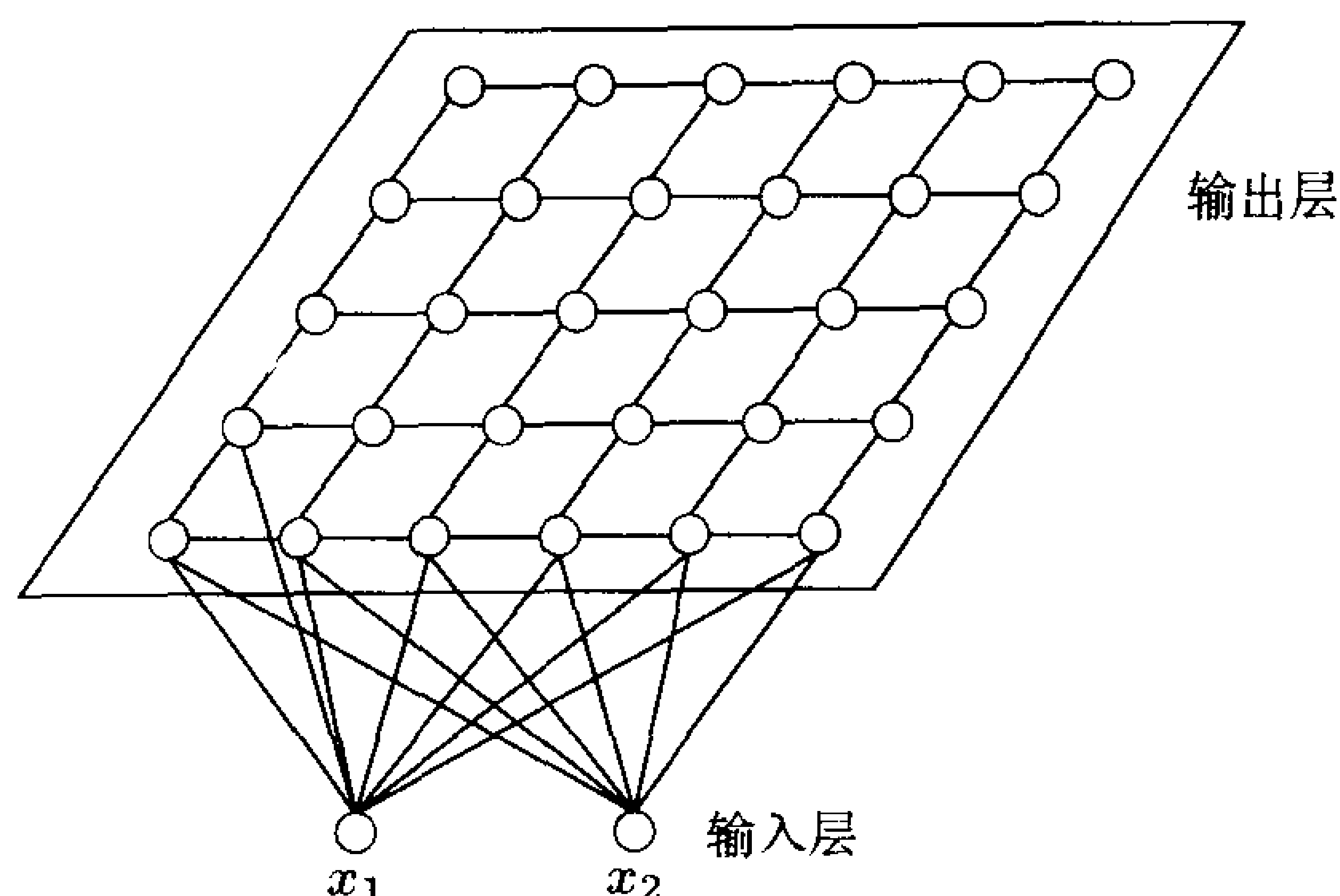


图 4.3.1　Kohonen 网络示意图

由 (4.3.20) 式可以看到, 当外界没有输入或输入与权的乘积之和比较小时, $\frac{\mathrm{d}y_j}{\mathrm{d}t}<0$, y_j 的值减小. 而当输入与权的乘积之和比较大时, $\frac{\mathrm{d}y_j}{\mathrm{d}t}$ 也就比较大, y_j 的值增长较快, 而 y_j 的增加又会引起 f 的增加, 从而使得 $\frac{\mathrm{d}y_j}{\mathrm{d}t}$ 减小. 最后系统收敛到稳定状态, 即 $\frac{\mathrm{d}y_j}{\mathrm{d}t}=0$, 此时 y_j 满足

$$y_j=f^{-1}\left(\sum_{i=1}^{n}w_{ij}x_i\right)\quad(1\leqslant j\leqslant m).\tag{4.3.21}$$

权系数 w_{ij} 的学习规则服从 Hebb 规则

$$\frac{\mathrm{d}w_{ij}}{\mathrm{d}t}=\alpha y_j x_i-\beta(y_j)w_{ij},\tag{4.3.22}$$

其中 α 是一个正的常数. (4.3.22) 式的第一项表示权系数的变化正比于输入 x_i 与输出 y_j 的乘积. 当 x_i 与 y_j 两个神经元同时处于兴奋状态时, 它们之间的连接应该得到加强, 相应的 w_{ij} 增长会加快, 称这种学习规则为 **Hebb 规则**. 这个规则与“条件反射”学说相一致, 并已得到神经细胞学说的证实. (4.3.22) 式的第二项是一个遗忘因子, 当外界没有输入时, 权系数 w_{ij} 将随着时间的推移而减小.

现在考虑一种特殊情形， y_j 为二值函数，取值为 $\{0,1\}$, 函数 $\beta(y_j)$ 满足

$$\begin{cases} \beta(y_j)=\alpha, & \text{当 } y_j=1 \text{ 时}, \\ \beta(y_j)=0, & \text{当 } y_i=0 \text{ 时}, \end{cases} \tag{4.3.23}$$

那么权系数的学习规则 (4.3.22) 式可改写为

$$\begin{cases} \dfrac{\mathrm{d}w_{ij}}{\mathrm{d}t}=\alpha(x_i-w_{ij}), & \text{当 } y_j=1 \text{ 时}, \\ \dfrac{\mathrm{d}w_{ij}}{\mathrm{d}t}=0, & \text{当 } y_j=0 \text{ 时}. \end{cases} \tag{4.3.24}$$

由此得到的结果是使学习后的权 w_{ij} 越来越接近输入值 x_i.

(4.3.24) 式中的 α 是与时间和距离都有关系的函数. 如果 $y_j=1$ 表示神经元兴奋， $y_j=0$ 表示神经元被抑制，那么由 (4.3.24) 式可以看到，权的调整是在兴奋的神经元 (即 $y_j=1$) 附近进行的. 比较靠近兴奋神经元的那些神经元，权调整系数 α 较大，而远离兴奋神经元的那些神经元，权调整系数 α 较小. α 与距离 r 之间的关系可表示为

$$\alpha(r)=a_1\mathrm{e}^{-\frac{r^2}{\sigma_1}}-a_2\mathrm{e}^{-\frac{r^2}{\sigma_2}}, \tag{4.3.25}$$

其中 r 为两个神经元之间的距离， $a_1, a_2, \sigma_1, \sigma_2$ 为常数，且 $a_1>a_2$, $\sigma_2>\sigma_1$. (4.3.25) 式所描述的函数形状像一顶墨西哥草帽 (见图 4.3.2).

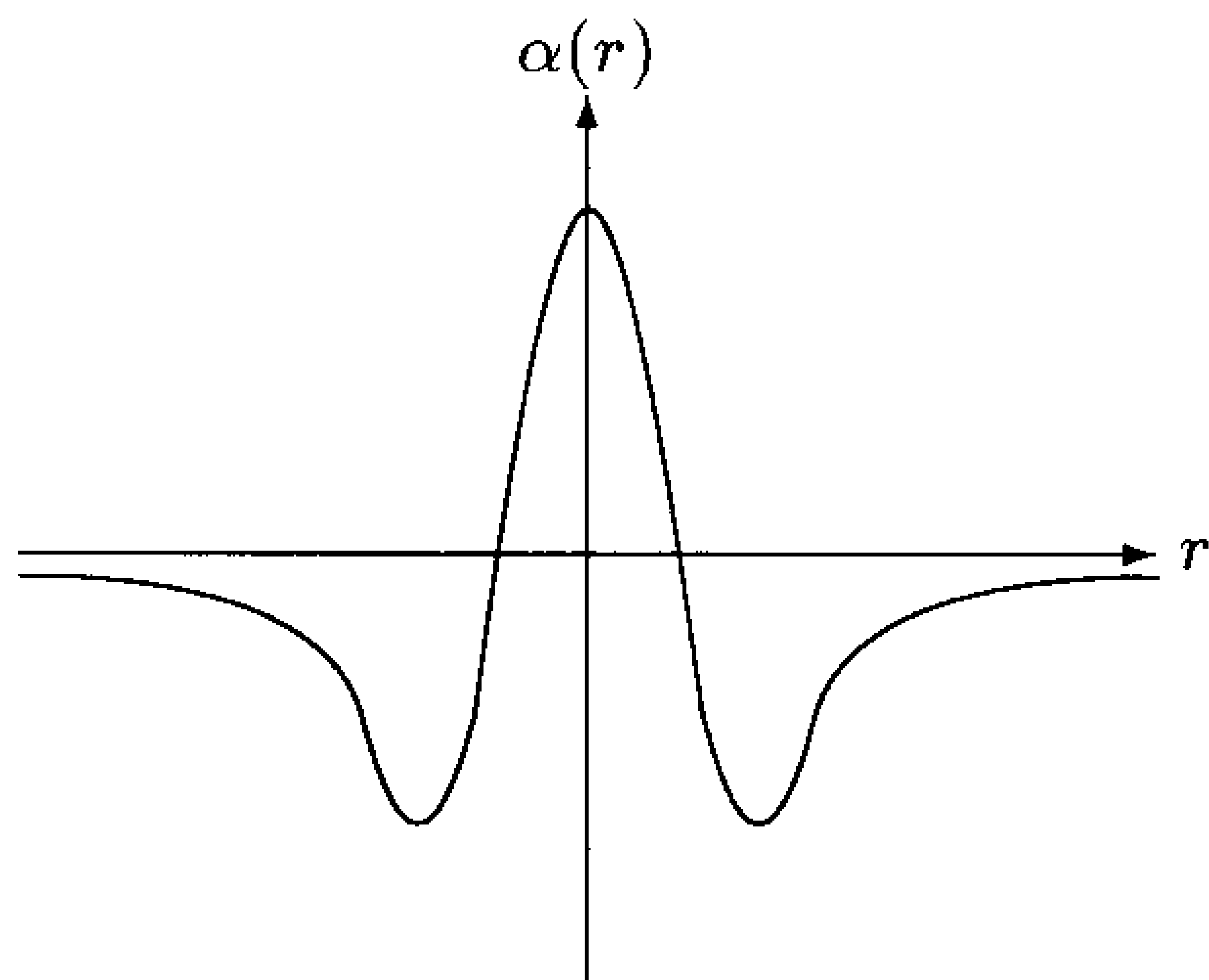


图 4.3.2

α 也可以用其它形式的函数表示，只要具有如下特点即可：离兴奋神经元越近，其值越大；离兴奋神经元越远，其值越小. 例如可以取火炉管函数 (见图 4.3.3)

$$\alpha(r)=\begin{cases}1, & r<R',\\ -b, & R'<r<R,\\ 0, & r>R\end{cases} \tag{4.3.26}$$

或礼帽函数 (见图 4.3.4)

$$\alpha(r)=\begin{cases}1, & r\leqslant R,\\ 0, & r>R\end{cases} \tag{4.3.27}$$

等等.

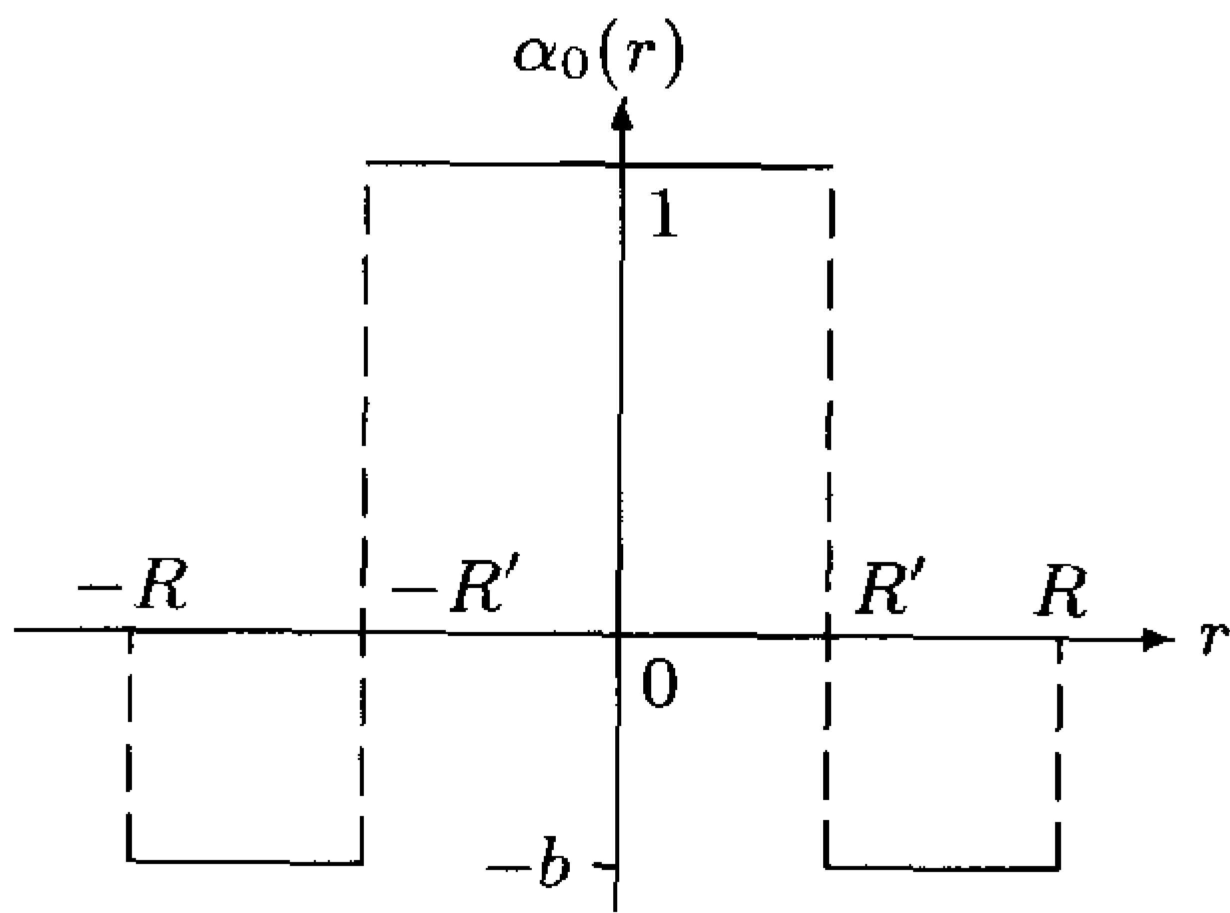


图 4.3.3

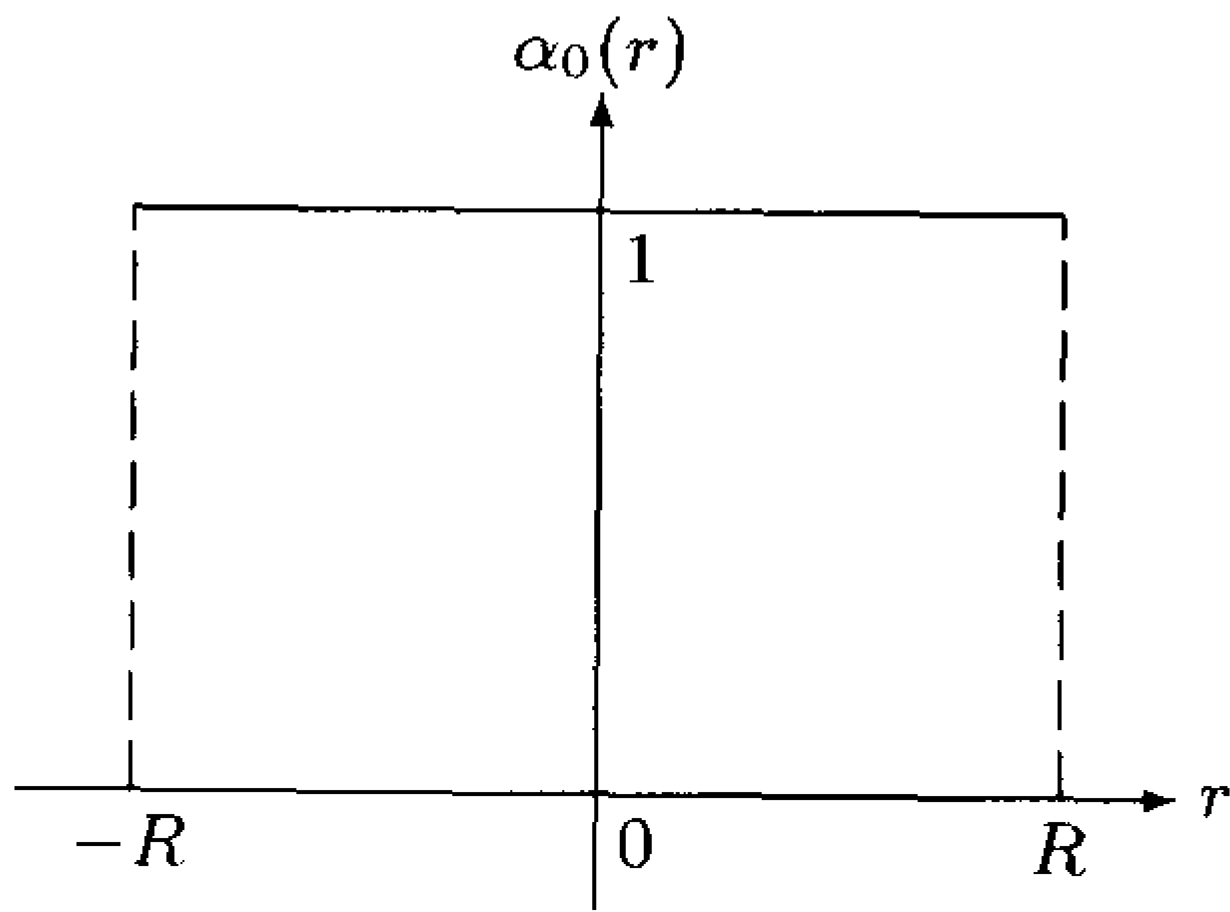


图 4.3.4

(4.3.24) 式中的 $\alpha(r)$ 又是时间 t 的函数，可以记为 $\alpha(r,t)$, 它是关于 t 的单调递减函数. α 与 t 之间的关系可以表示为

$$\alpha(t) = A_1 e^{-\frac{t}{C_1}} \tag{4.3.28}$$

或

$$\alpha(t) = A_2 \left(1 - \frac{t}{C_2}\right), \tag{4.3.29}$$

其中 A_1, C_1, A_2, C_2 均为常数.

对于兴奋的那个神经元 j, 在其周围区域 D_j 内的神经元得到不同程度的兴奋，而在区域 D_j 以外的神经元则被抑制. 区域 D_j 可以取不同的形状，如正方形 (见图 4.3.5) 或六边形 (见图 4.3.6) 等等.

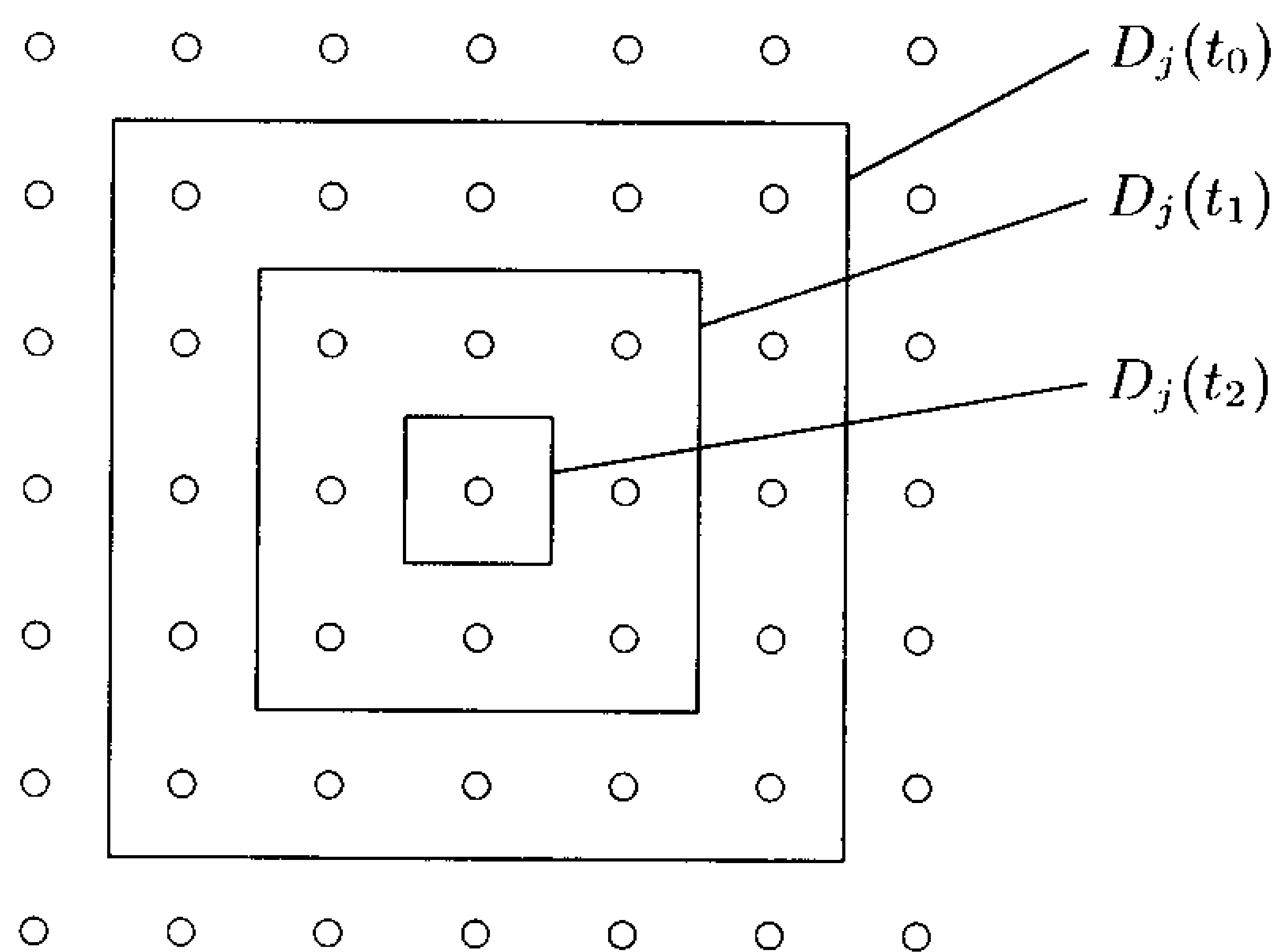


图 4.3.5　$D_j(t)$ 取四边形的情形

区域 D_j 还是时间 t 的函数，随着时间的变化而逐步减小，直到只剩下一个神经元为止，因此 D_j 可表示为

$$D_j(t) = A_1 + A_2 e^{-\frac{t}{C}}, \tag{4.3.30}$$

其中 A_1, A_2, C 均为常数.

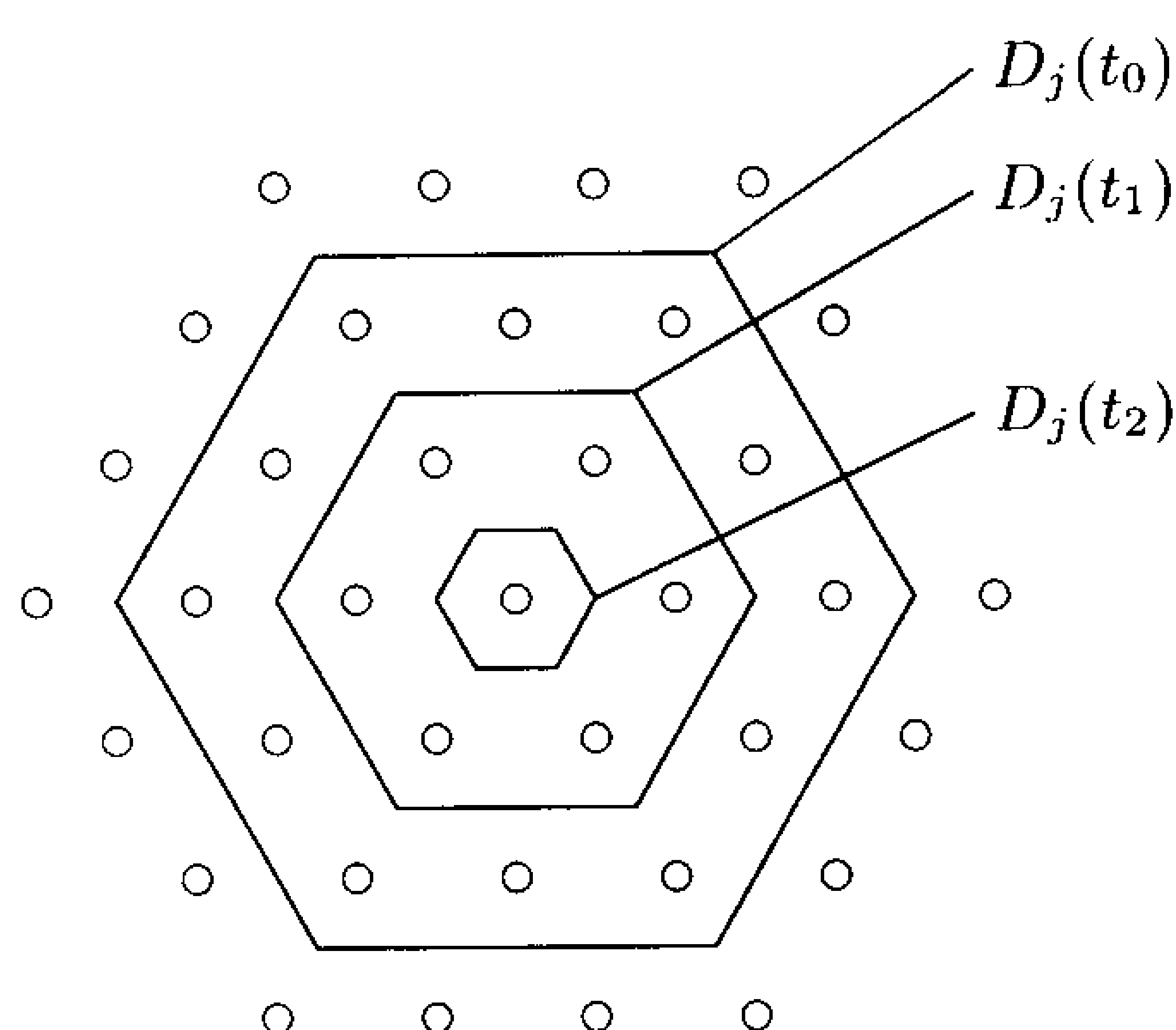


图 4.3.6　$D_j(t)$ 取六边形的情形

在 TSP 问题中，如果把输入样本看作是一个城市的集合，其中每一个城市是定位在二维平面上的，如同地图上表示城市的位置坐标一样，在集合 C 中每个样本用 $C_j(x_j, y_j)$ 来表示，因此这种网络的输入为两个神经元，表示城市的 x, y 坐标，这种网络的输出为 100 个以上的神经元，组成了一个平面. 网络中的连接权初始时为一个随机数，按照前面的算法进行学习，其邻域 D_j 一开始取得很大，而随时间减小，$\alpha(t, r)$ 也随之改变. 由于 Kohonen 网络中的映照特性和聚类特性，使得在地图上相近的城市在输出平面中的位置也比较相近，把相近的点连成链状，这个链无疑是按照城市远近排列的，所以这个链就是 TSP 问题的解.

4.3.3　TSP 问题 —— 遗传算法

TSP 问题还可以用遗传算法来解决.

遗传算法 是一种建立在生物遗传基础上的算法，它利用基因组成各种基因组，对基因组进行遗传、变异和交叉，得到下一代基因组. 在这个繁殖过程中，那些对应于符合要求的基因组获得更多的生存机会. 经过反复繁殖和进化得到一个较优的基因组，这个基因组表现出的特性与要求的特性一致.

我们结合 TSP 问题具体地介绍遗传算法的步骤.

设有 n 个城市, 用 d_{ij} 表示第 i 个城市与第 j 个城市之间的距离. 因为 $d_{ij}=d_{ji}$, n 个城市有 $\frac{n(n-1)}{2}$ 个距离. 于是我们可以得到 TSP 问题的遗传算法的具体步骤:

(1) 计算 d_{ij} $(i,j=1,2,\cdots,n)$ 并随机组合成基因组, 每个基因组由 n 个基因构成, 表示一条推销员的路径;

(2) 计算每个基因组的总距离 $d=d_{i_1i_2}+d_{i_2i_3}+\cdots+d_{i_{n-1}i_n}+d_{i_ni_1}$;

(3) 如果问题只要求距离 d 小于一个给定值, 记为 $\widetilde{d}$, 那么只要 $d<\widetilde{d}$, 算法就结束; 否则进入选择阶段;

(4) 在选择阶段中, 对于路程最优的那些基因组进行复制, 对于坏的基因进行删除, 这样留下的基因组得到的路径都比较优;

(5) 进行变异, 或突变, 即可以从上一代突变到子代, 突变的方法是:

(i) 可以将某个基因组代表的路径中城市的次序进行颠倒, 每次只颠倒一部分;

(ii) 对换两个城市之间的位置;

(iii) 可以用两个基因组合成 n 个基因组的办法.

所有的突变方法因为改动不大, 所以计算量也不大, 而产生出新的一代比上代多. 对这些子代再进行步骤 (2) 的计算, 这样新的一轮进化又开始了, 因此子孙的增加使输入的样本面逐步扩大. 但是, 十分明显, 计算并不因为迭代的次数增加而增大, 因为在程序中对子孙代进行了删除、优化, 因此每次迭代的计算量是固定的, 最后可得到要求的较优解.

值得注意的是, 遗传算法一般不能得到最优解, 而只能得到一个近似最优解, 即对 TSP 问题只能得到近似最短路径.

从上述算法可以看到, 遗传算法只取一部分的基因组作为初值, 然后沿着进化方向繁殖子代, 而每个子代必是上面所有排列中的一个, 计算量大为减少, 从而速度也比较快.

§4.4 金融工程

股市扩容

中国大陆的证券市场是被改革的浪潮重新推上大陆的经济舞台的, 在短短的几年中其发展速度引起世人注目, 现代化手段的应用使我们的交易在技术上接近国际上有百余年发展史的著名西方证券交易机构. 然而我们毕竟是一个新兴的尚未成熟的股市, 由于管理经验的不足、法制的不够完善、股民的复杂心态等诸多原因, 股市经历了跌荡起伏的过程, 付出了高昂的代价. 上海股市在其发展历程中, 虽没有出现 "黑色星期一" 那样震惊世界的波澜, 但其大起大伏的频繁也证实了这一点. 每次代价的付出, 使得股民理智了一些, 股市也向成熟迈出了一步. 从上海股市几年来异常波动的成因分析可看出, 股市受外来因素影响要强于内在因素的影响, 但从时间进程看, 市场的内在机制正越来越发挥自动调节作用, 而外来因素影响在逐步缩小, 这就使得我们理性地根据内在机制来分析股市成为必须和可能.

我国的宏观经济形势总体趋好, 正步入 "绿灯区". 经济的稳步增长、物价涨幅的进一步下降、利率的下调, 无疑对资金的入市产生重要影响, 有力地支持了证券市场的发展. 而在另一方面, 证券市场在整个国民经济发展中举足轻重的地位也被普遍认识. 国务院领导也提出稳步发展、适当加快的方针, 明确了发展证券市场的主基调. 在推出活跃市场种种举措的同时, 证券市场本身也将加速扩容以迎接新的高潮.

几年来股市变化的实际表明扩容一方面给股市带来了生机和规模, 另一方面也带来了震荡, 这对于像我国大陆股市那样的新兴市场是必然的, 从本质上来说正是股市资金量的平衡 — 变化 — 平衡的动态变化的表现. 我们根据这一规律建立起资金流的平衡模型来研究股市扩容问题, 为领导层决策时参考.

股市的供求关系是把握扩容的关键, 股市的流通市值表征了股市的资金量, 我们据此来建立模型.

记 S_{ij} 表示 i 时刻第 j 种股票的流通面值，h_{ij} 表示相应的股价，那么 i 时刻总的流通市值 V_i 可表示为

$$V_i = \sum_j h_{ij} S_{ij}. \tag{4.4.1}$$

如果将上述 S_{ij} 及 h_{ij} 中的 j 去掉，以表示所有的股票，而不特指第 j 种股票，即改记 S_i 为 i 时刻所有股票的总的流通面值，h_i 为所有股票的平均股价，那么

$$V_i = h_i S_i. \tag{4.4.2}$$

股价的变化、股市的扩容都会引起股市中资金量的变化，但是总的资金应处于动态平衡的状态，亦即它应是守恒的. 因此在一定的单位时间内

$$\Delta V_i = \mathrm{In}_i - \mathrm{Out}_i, \tag{4.4.3}$$

这里 In_i 表示资金的入市量，而 Out_i 表示资金的出市量. 从影响股市的因素来说，也可以表示为

$$\mathrm{In}_i - \mathrm{Out}_i = F_{1i} + F_{2i}. \tag{4.4.4}$$

F_{1i} 表示股市内在因素产生的对资金量的影响，而 F_{2i} 则是外在因素产生的对资金量的影响. 这样，对于股市资金量的动态平衡有

$$\Delta(h_i S_i) = F_{1i} + F_{2i}, \tag{4.4.5}$$

相应的微分形式是

$$\mathrm{d}(hS) = F_1 + F_2. \tag{4.4.6}$$

单位时间内的股市资金量动态平衡方程 (4.4.5) 可展开为

$$h_i \Delta S_i + S_i \Delta h_i + \Delta S_i \Delta h_i = F_{1i} + F_{2i}. \tag{4.4.7}$$

(4.4.7) 式中左端各项可以作如下描述：$h_i \Delta S_i$ 表示平均股价不变时，单由扩容 ΔS_i 而引起的资金量的增加；$S_i \Delta h_i$ 表示在不扩容

情况下，平均股价变化 Δh_i 而引起的资金量的增加；$\Delta S_i \Delta h_i$ 则表示扩容部分其股价变化量而引起的资金量的增加.

为了更直接反映出扩容和上证指数之间的关系，对 1994~1995 年的平均股价和上证指数作了统计分析，发现它们具有极强的线性相关性. 以 H 表示上证指数，有

$$\begin{cases} h = aH + b, \\ a = 0.0077, \quad b = 0.1942, \quad r = 0.95. \end{cases} \tag{4.4.8}$$

于是 (4.4.7) 式可改写成

$$(aH_i + b)\Delta S_i + S_i a \Delta H_i + \Delta S_i a \Delta H_i = F_{1i} + F_{2i} \tag{4.4.9}$$

或写成

$$(aH_{i+1} + b)\Delta S_i + S_i a \Delta H_i = F_{1i} + F_{2i}. \tag{4.4.10}$$

从 (4.4.10) 式出发作如下定性分析：

(1) 股市资金量不变的情况

此时 $F_{1i} + F_{2i} = \Delta V = 0$, 即既无外在因素又无内在因素使得资金入市或出市，股民观望，股市盘整. (4.4.10) 式简化成

$$(aH_{i+1} + b)\Delta S_i + S_i a \Delta H_i = 0, \tag{4.4.11}$$

所以

$$\Delta H_i = -\frac{(aH_{i+1} + b)}{S_i a}\Delta S_i. \tag{4.4.12}$$

如果此时不扩容，$\Delta S_i = 0$, 则 $\Delta H_i = 0$, 意味着指数仅有小的波动.

如果此时扩容，$\Delta S_i > 0$, 则 $\Delta H_i < 0$, 指数将要发生下降.

(2) 不扩容的情况

此时 $\Delta S_i = 0$, (4.4.10) 式简化为

$$S_i a \Delta H_i = F_{1i} + F_{2i} \tag{4.4.13}$$

或写成

$$\Delta H_i = \frac{F_{1i} + F_{2i}}{S_i a}. \tag{4.4.14}$$

如果资金入市量和出市量不变，指数将发生变化，其变化量 ΔH_i 与股市的容量成反比，容量大时指数的变化 $|\Delta H_i|$ 相应要小，反之则大.

如果资金入市，总量增大，$F_{1i} + F_{2i} > 0$，则 $\Delta H_i > 0$，指数将上扬；如果资金出市，总量减小，$F_{1i} + F_{2i} < 0$，则 $\Delta H_i < 0$，指数将下降.

(3) 正常扩容的情况

此时有 $\Delta S_i > 0$，(4.4.10) 式可写成

$$\Delta H_i = \frac{F_{1i} + F_{2i} - \Delta S_i(aH_{i+1} + b)}{S_i a}. \tag{4.4.15}$$

如果资金总量不变或减少，即 $F_{1i} + F_{2i} \leqslant 0$，则必然有指数下降.

如果资金总量有所增加，$F_{1i} + F_{2i} > 0$，但不足以保证 $F_{1i} + F_{2i} \geqslant \Delta S_i(aH_{i+1} + b)$，指数依然是要下降的.

如果形势很好，有充分的入市量，使得 $F_{1i}+F_{2i} > \Delta S_i(aH_{i+1}+b)$，那么指数必将上升.

以上定性结论和几年来股市变化基本上是吻合的.

很显然，股市资金总量的变化是由资金入市量和出市量所决定的. 影响资金的入市和出市的因素有两类：一类是市场内在机制因素 (F_1)，如工业总产值、消费品零售价、价格指数、成交量、储蓄存款、利率等；另一类是外在因素 (F_2)，主要是政策、扩容、消息等，同时还包括一些不可预见的因素.

如何在方程 (4.4.5) 和 (4.4.6) 中来确定 F_1 和 F_2 的表达式，无论是线性的或非线性的形式. 由前面的分析，影响资金的入市和出市的因素非常复杂，很难用一个明确的解析表达式来表示. 为此，我们应用 BP 神经网络来求解.

将影响因素作为输入向量，记为 $x = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，股市资金的入市量作为输出变量，记为 y. 输入向量与输出变量之间的关

系满足

$$y=f\left(\sum_{i=1}^{n}w_ix_i\right), \tag{4.4.16}$$

其中 w_i $(1 \leqslant i \leqslant n)$ 为权系数， $f(u)$ 为单调增加的非线性函数. 通过 BP 神经网络的学习，可以得到一组最佳权向量 $w^*=(w_1^*,w_2^*,\cdots,w_n^*)^{\mathrm{T}}$. 这样，利用这组最佳权向量 w^*, 得到资金入市量与各种影响因素之间的关系，结果记为 F^*.

对于所建立的模型，从 (4.4.10) 式可导出

$$\Delta S_i=\frac{F_{1i}+F_{2i}-S_ia\Delta H_i}{aH_{i+1}+b}, \tag{4.4.17}$$

这就是扩容速度的表达式. 可以看到，影响扩容速度的因素有 (i) 资金入市和出市的总量 $F_{1i}+F_{2i}$; (ii) 流通面值 S_i; (iii) 指数 H_{i+1}; (iv) 指数的波动 ΔH_i. 我们的目的就是希望通过确定的指数 H_i 和允许的波动 ΔH_i, 来确定 ΔS_i 的大小，其中关键的困难之处就是资金入市量和出市量的问题. 我们结合 F^* 的实证分析结果，研究股市扩容的速度问题，得到了满意的结果.

§4.5　分　　类

4.5.1　样本集的分类

如图 4.5.1 所示，在二维平面中有一个由两类样本构成的样本集. 一类为 A 类样本，集中在平面的左上角，用 “∘” 表示；另一类为 B 类样本，集中在平面的右下角，用 “×” 表示. 我们希望找到一根直线

$$w_1x_1+w_2x_2-\theta=0, \tag{4.5.1}$$

把 A, B 两类样本分隔开来.

如果 A, B 两类样本是线性可分的，即可用一根直线将两类样本分隔开来，而且有一段距离 (参见图 4.5.1), 那么形如 (4.5.1) 式的解有无数个. 我们采用单层感知器 (perceptron) 来求解这个问题.

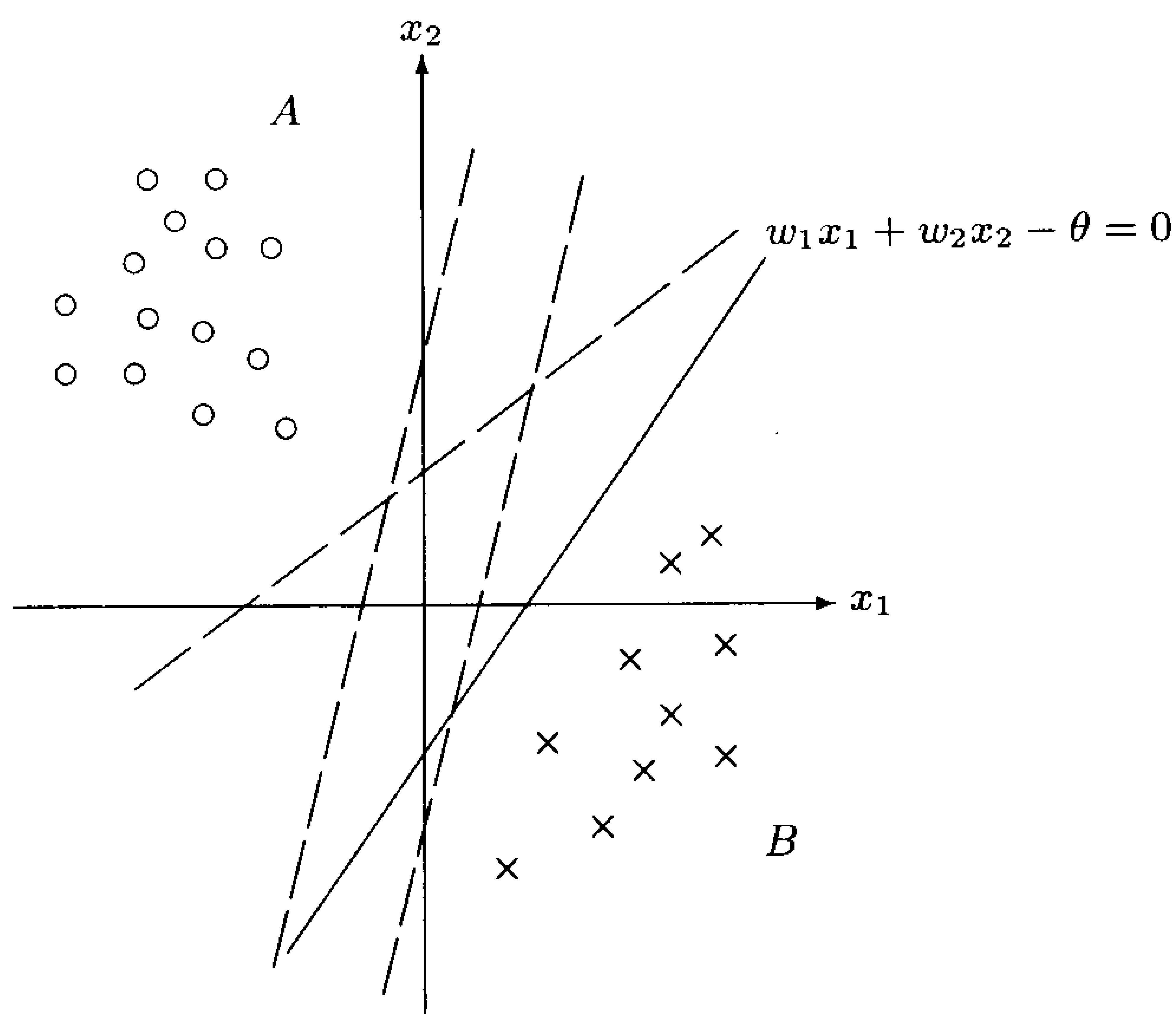


图 4.5.1　样本集的线性划分

感知器是由 F. Rosenblatt 于 1957 年提出的. 设 $x=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 为输入向量， y 为输出量，输入与输出之间满足

$$y=f\left(\sum_{i=1}^{n}w_ix_i-\theta\right), \tag{4.5.2}$$

其中 w_i $(1\leqslant i\leqslant n)$ 为权系数， θ 为阈值，函数 $f(u)$ 为分段常数函数

$$f(u)=\begin{cases}1, & u\geqslant 0,\\ -1, & u<0.\end{cases} \tag{4.5.3}$$

令 $w_0=\theta$, $x_0=-1$, 那么

$$y=f\left(\sum_{i=0}^{n}w_ix_i\right). \tag{4.5.4}$$

于是问题化为已知两个样本集分别为 $\{x^A\}$, $\{x^B\}$, 要求权系数和

阈值 w_i $(0 \leqslant i \leqslant n)$, 使得

$$y = f\left(\sum_{i=0}^{n} w_i x_i^A\right) = 1, \tag{4.5.5}$$

$$y = f\left(\sum_{i=0}^{n} w_i x_i^B\right) = -1. \tag{4.5.6}$$

单层感知器的学习过程是一种有教师的学习过程. 设教师 t 满足

$$t = \begin{cases} 1, & \text{当输入 } x^A \text{ 时}, \\ -1, & \text{当输入 } x^B \text{ 时}, \end{cases} \tag{4.5.7}$$

那么整个学习过程为

(1) 随机给出权和阈值的初值 w_i^0 $(0 \leqslant i \leqslant n)$, 初值为非零的小数, 即 $0 < |w_i^0| \ll 1$;

(2) 任选样本集中的一个样本 x 作为输入向量, 计算实际输出

$$y = f\left(\sum_{i=0}^{n} w_i x_i\right); \tag{4.5.8}$$

(3) 若 $x \in x^A$, 令 $t = 1$; 若 $x \in x^B$, 令 $t = -1$;

(4) 按下式调节权和阈值

$$w_i(k+1) = w_i(k) + \eta(t - y(k))x_i(k), \tag{4.5.9}$$

其中 k 为迭代次数, $0 < \eta \leqslant 1$, 用于控制调节速度. 通常 η 不能取得太小, 否则会使 w_i 的收敛速度太慢; 但 η 也不能取得太大, 否则会影响 w_i 的稳定;

(5) 在 x^A 或 x^B 集中选取另一个样本进行学习, 即重复 (2)~(4), 直到权系数 w 对一切样本均稳定不变为止, 即 $w_i(k+1) = w_i(k)$ $(i = 0, 1, 2, \cdots, n)$.

为了防止权调整的不均匀性, 学习时样本的选取最好在 x^A 与 x^B 两个样本集中轮流进行.

例如, 设输入样本 $x = (x_1, x_2)^{\mathrm{T}}$ 与教师 t 由表 4.5.1 给出.

表 4.5.1

输入样本 x		教师 t
x_1	x_2	
0	0	0
0	1	0
1	0	0
1	1	1

由单层感知器计算可得 $w_0 = 1.5$, $w_1 = 1$, $w_2 = 1$, 于是分隔直线为

$$x_1 + x_2 - 1.5 = 0, \tag{4.5.10}$$

而输出函数为

$$y = f(x_1 + x_2 - 1.5), \tag{4.5.11}$$

其中

$$f(u) = \begin{cases} 1, & u \geqslant 0, \\ 0, & u < 0. \end{cases} \tag{4.5.12}$$

这个例子说明了“逻辑与”运算可以用单层感知器来实现. 若将“逻辑与”的输入样本看成是直角坐标系下单位正方形的 4 个顶点 (参见图 4.5.2), 那么直线 (4.5.10) 将这 4 个顶点分成两类, 右上

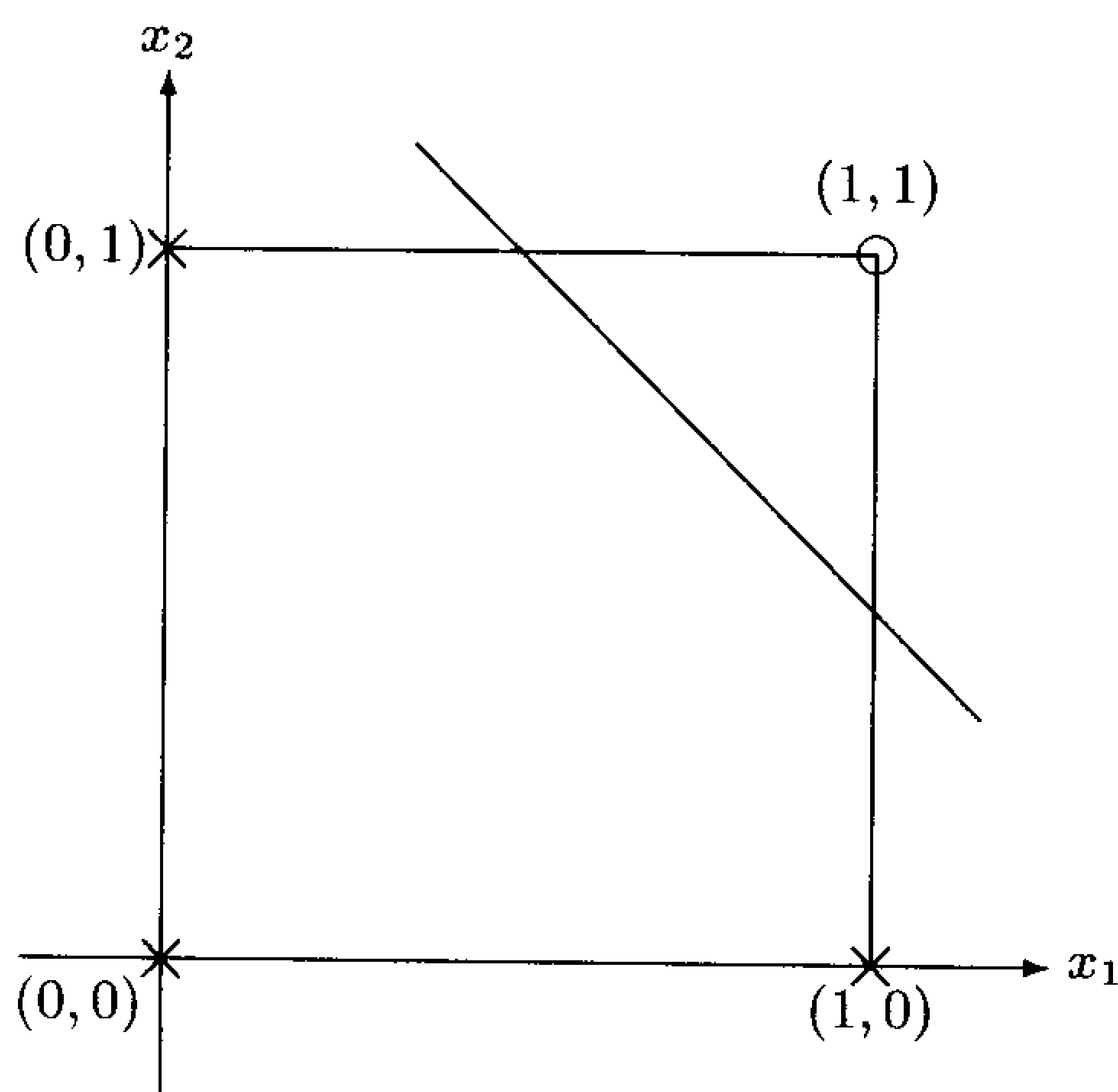


图 4.5.2 用单层感知器实现“逻辑与”运算

方为 A 类，输出为 1, 左下方为 B 类，输出为 0.

同样地，“逻辑或”运算也可以用单层感知器来实现. 设输入样本 $x=(x_1,x_2)^{\mathrm{T}}$ 与教师 t 由表 4.5.2 给出.

表 4.5.2

输入样本 x		教师 t
x_1	x_2	
0	0	0
0	1	1
1	0	1
1	1	1

由单层感知器计算可得 $w_0=0.5$, $w_1=1$, $w_2=1$, 于是分隔直线为

$$x_1+x_2-0.5=0, \tag{4.5.13}$$

而输出函数为

$$y=f(x_1+x_2-0.5), \tag{4.5.14}$$

其中 $f(u)$ 由 (4.5.12) 给出. “逻辑或”运算的几何图示见图 4.5.3.

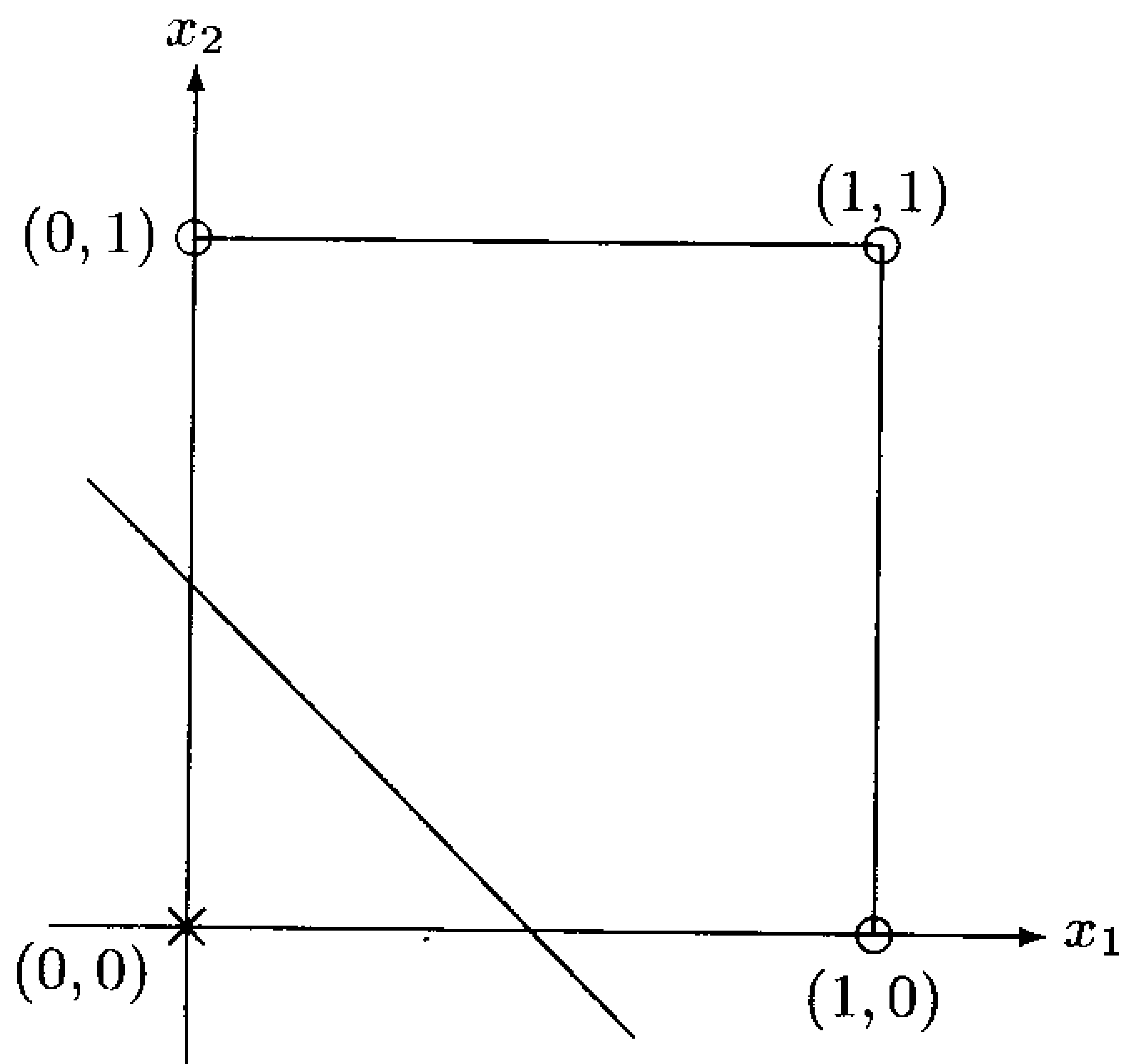


图 4.5.3 用单层感知器实现“逻辑或”运算

但是单层感知器只能满足线性分类，如果两类样本不能用一个超平面分开，那么单层感知器就无能为力了.

例如“逻辑异或”运算就无法用单层感知器实现. 事实上，设输入样本 $x=(x_1,x_2)^{\mathrm{T}}$ 与教师 t 由表 4.5.3 给出.

表 4.5.3

输入样本 x		教师 t
x_1	x_2	
0	0	0
0	1	1
1	0	1
1	1	0

由单层感知器应成立

$$\begin{aligned}
0\cdot w_1+0\cdot w_2-\theta&<0,\\
0\cdot w_1+1\cdot w_2-\theta&\geqslant 0,\\
1\cdot w_1+0\cdot w_2-\theta&\geqslant 0,\\
1\cdot w_1+1\cdot w_2-\theta&<0.
\end{aligned}$$

显然上述不等式方程组无解，即单层感知器无法解决“逻辑异或”问题. 从几何图形来看， A 类顶点 $(0,1),(1,0)$ 与 B 类顶点 $(0,0)$, $(1,1)$ 分别在单位正方形的对角顶点上，无法用一根直线将它们分开 (参见图 4.5.4).

又如图 4.5.5 所示的二维平面中， A 类样本分布在原点的附近， B 类样本分布在 A 类的外部区域中，两类样本不能用直线分隔开来.

对于图 4.5.5 的问题，就是寻找一个区域，使其内部为 A 类样本，其外部为 B 类样本. 可以在二维输入空间中划出 3 根直线，因为它们的权系数 w_i 和阈值 θ 各不相同，因此 3 根直线的斜率与

截距也不相同 (如图 4.5.6 所示). 将这 3 个单元所得到的直线作相应的逻辑运算

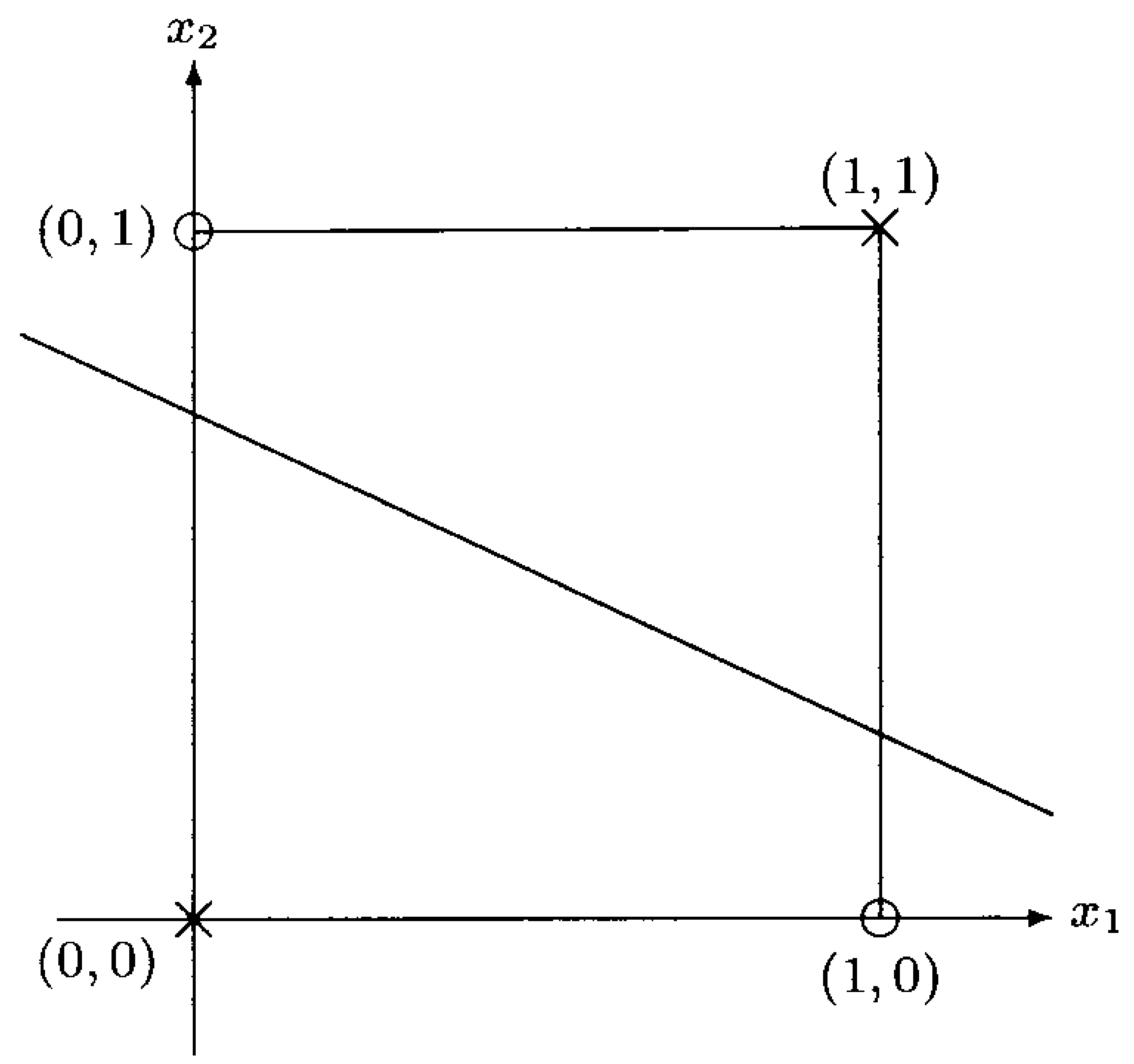


图 4.5.4 "逻辑异或" 运算

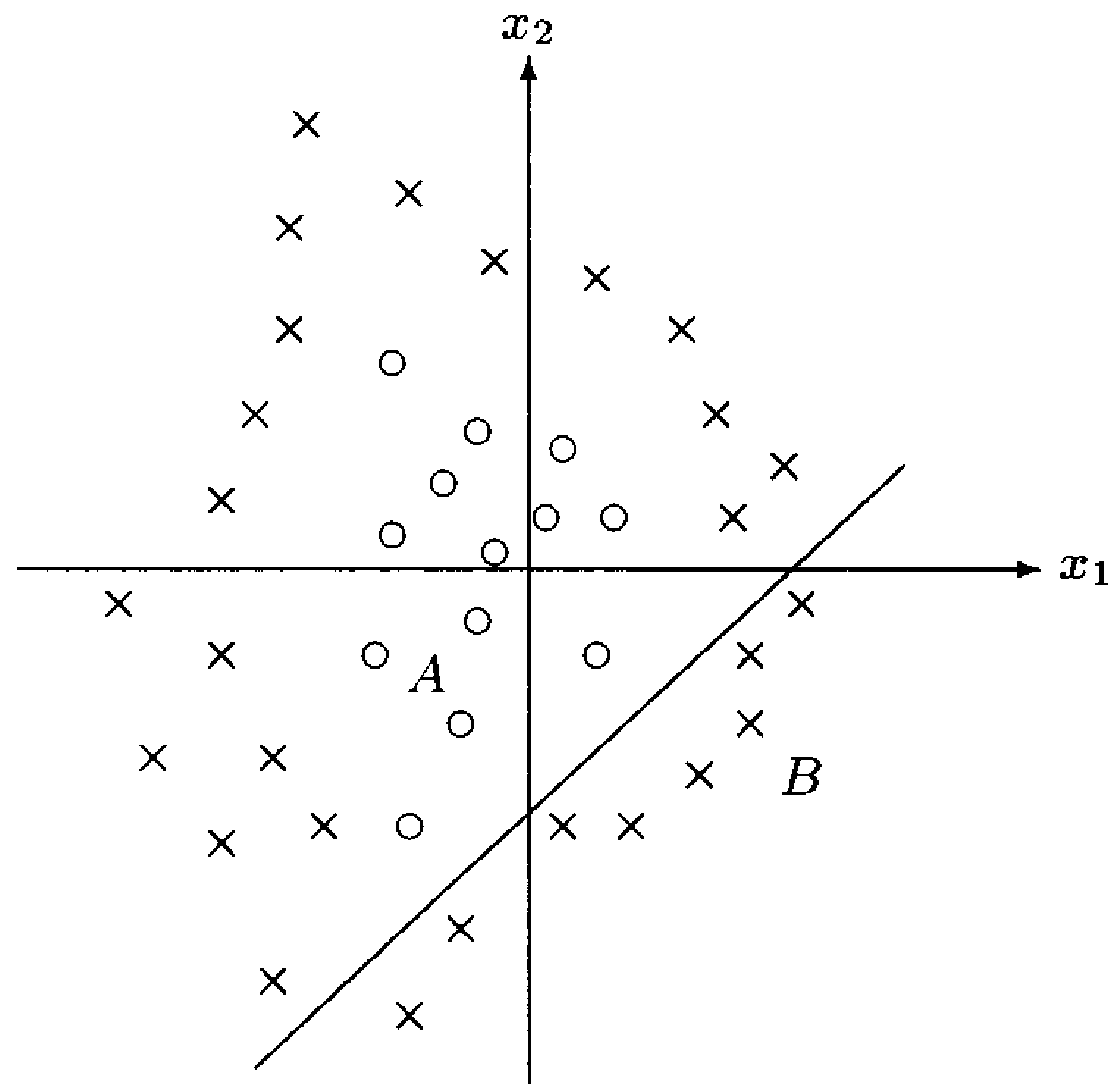


图 4.5.5 不能线性划分的样本集

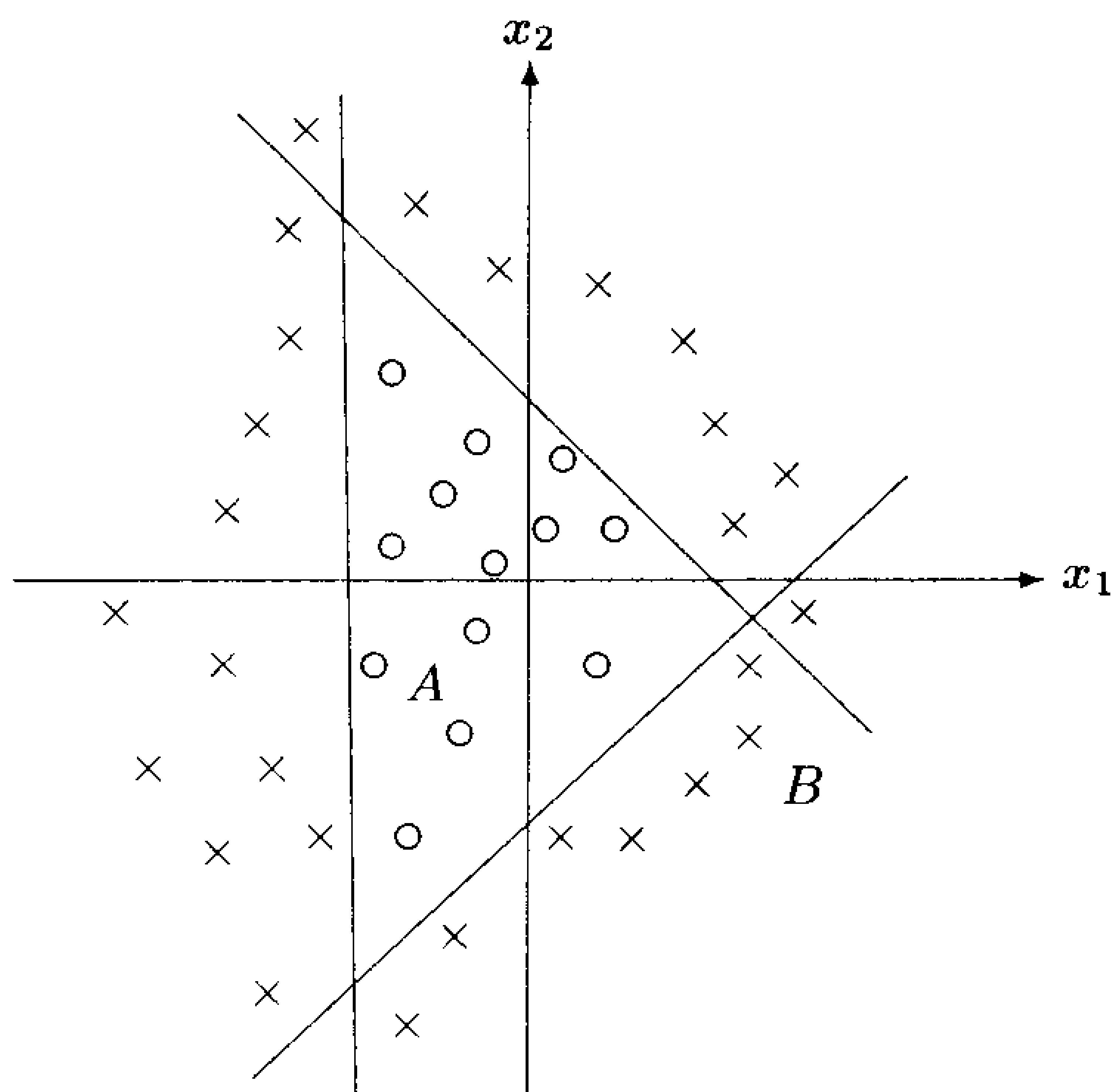


图 4.5.6　多层感知器网络对样本集的划分

$$
\begin{aligned}
z = \{(x_1, x_2) \mid & (w_{11}x_1 + w_{21}x_2 - \theta_1 > 0) \\
& \cap (w_{12}x_1 + w_{22}x_2 - \theta_2 > 0) \\
& \cap (w_{13}x_1 + w_{23}x_2 - \theta_3 > 0)\},
\end{aligned} \tag{4.5.15}
$$

其中 z 为输出单元. 这样得到一个封闭区域即可正确划分两类样本.

显然, 隐单元到输出单元之间为“逻辑与”的关系, 它可以用感知器实现, 于是, 图 4.5.5 的问题可以用 3 层感知器来完成.

对于“逻辑异或”运算可以抽象为图 4.5.7 的问题.

这个问题可以分别采用 4 层或 3 层感知器网络来完成.

若用 4 层感知器网络来实现, 第 1 层为输入层, 第 2 层起线性划分作用, 第 3 层为“逻辑与”运算, 第 4 层为“逻辑或”运算. 划分的区域为

$$
\begin{aligned}
z = \left\{ (x_1, x_2) \middle| \left(\bigcap_{j=1}^{3} (w_{1j}x_1 + w_{2j}x_2 - \theta_j > 0) \right) \right. \\
\left. \bigcup \left(\bigcap_{j=4}^{6} (w_{1j}x_1 + w_{2j}x_2 - \theta_j > 0) \right) \right\}.
\end{aligned} \tag{4.5.16}
$$

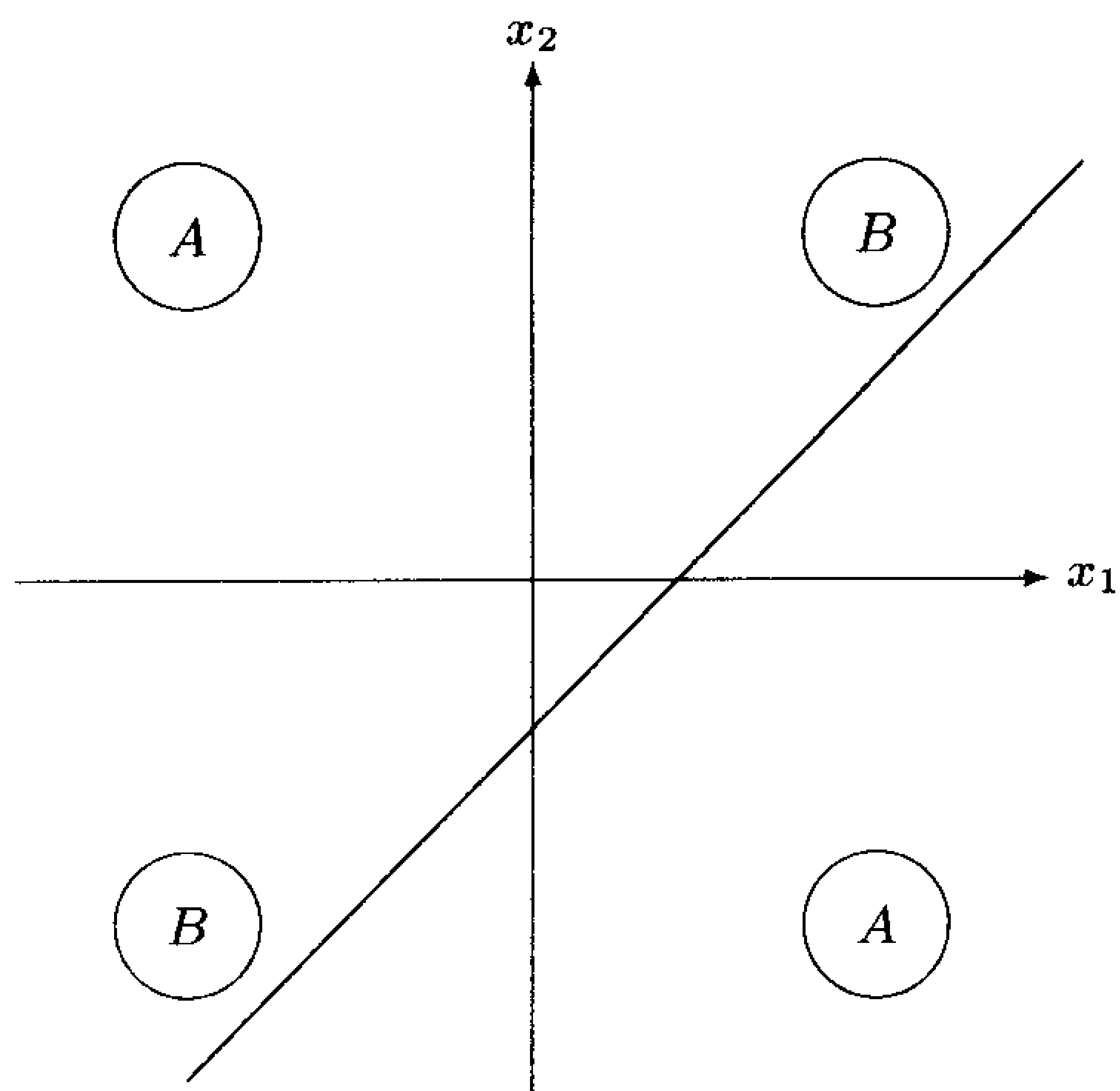


图 4.5.7　另一类不能线性划分的样本集

这里有 6 根直线，各用 3 根直线组成了两个三角形的区域，这是用“逻辑与”来完成的．两个区域内部都属于 A 类，这又是“逻辑或”的关系，其它区域则为 B 类 (见图 4.5.8).

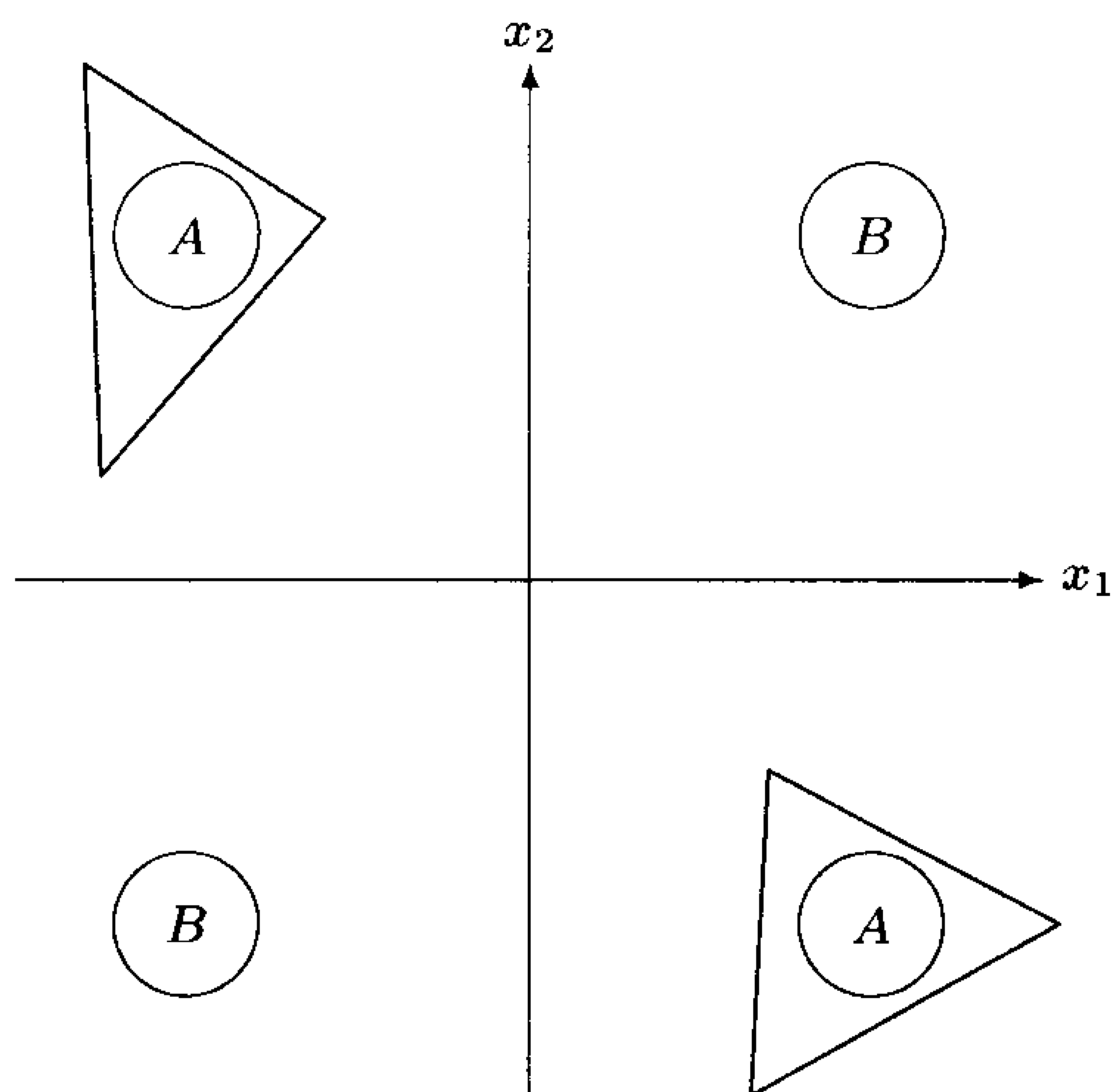


图 4.5.8　用 4 层感知器网络对样本集进行划分

若用 3 层感知器网络来完成，第 1 层为输入层，第 2 层起线性划分作用，第 3 层为“逻辑与”运算. 划分的区域为

$$z = \{(x_1, x_2) \mid (w_{11}x_1 + w_{21}x_2 - \theta_1 > 0) \bigcap (w_{12}x_1 + w_{22}x_2 - \theta_2 > 0)\}. \quad (4.5.17)$$

这里用两根直线的“逻辑与”运算完成了样本的划分 (见图 4.5.9).

可以看出，多层感知器可通过单层感知器进行适当的组合达到任何形状的划分.

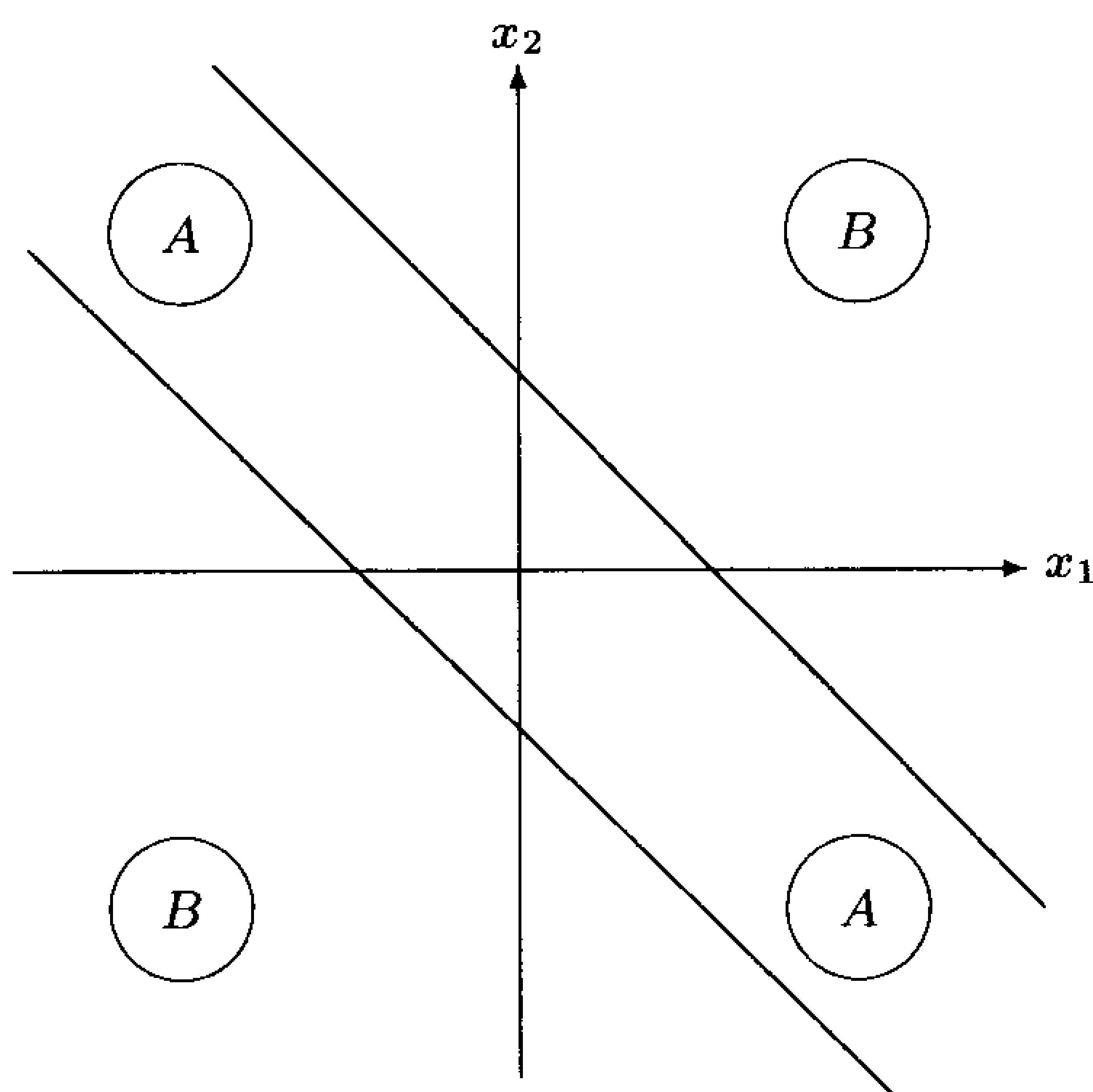


图 4.5.9　用 3 层感知器网络对样本集进行划分

4.5.2　T-C 模式分类问题

T-C 分类是一个典型的分类问题. 设输入一串字符序列，序列中的字符由“T”和“C”两个字母构成，每个字母由 3×3 个像素组成，可分别旋转 $0°, 90°, 180°, 270°$ (见图 4.5.10)，要求能够区分出一组输入序列中的每个字母. 我们用 3 层 BP 神经网络来完成这一工作. 设 BP 网络的输入是一个 3×3 的矩阵 $x = (x_{ij})$，其中

的元素 $x_{ij}=1$ 表示该点是黑色的， $x_{ij}=0$ 表示该点是白色的. 例如字母 "T" 在不旋转时的输入矩阵为

$$x=\begin{pmatrix}1&1&1\\0&1&0\\0&1&0\end{pmatrix}, \tag{4.5.18}$$

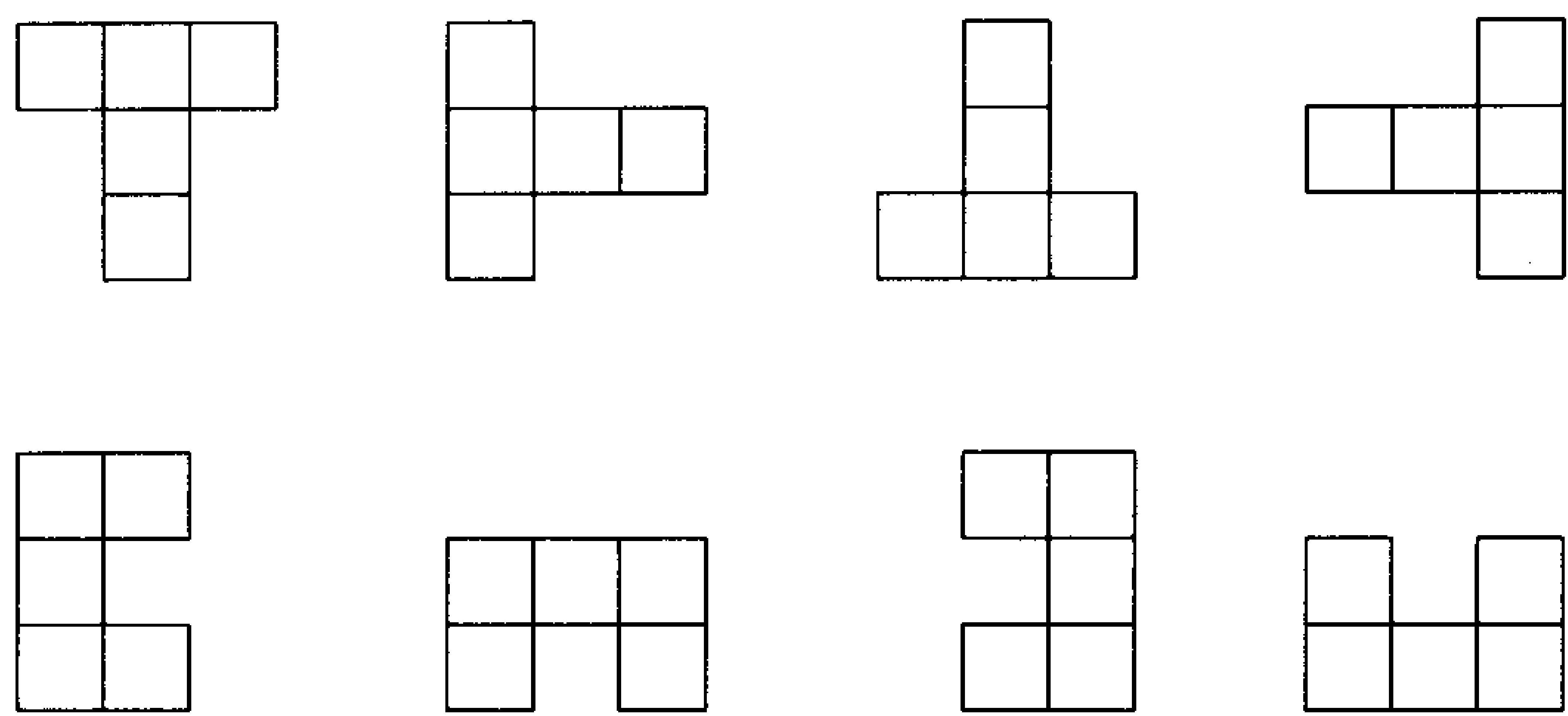


图 4.5.10　"T" 和 "C" 两个字母的旋转图形

字母 "C" 在不旋转时的输入矩阵为

$$x=\begin{pmatrix}1&1&0\\1&0&0\\1&1&0\end{pmatrix}, \tag{4.5.19}$$

等等.

中间层的每个隐单元并不与所有的输入单元相连接，而只与 3×3 个输入单元相连接，对每个隐单元，其连接权都是相同的. 为了使隐单元能覆盖输入的整个平面，隐单元数与输入单元数几乎相等，如输入单元数为 n^2, 而隐单元数为 $(n-1)^2$, 输出单元数为 1, 输出单元与每个隐单元相连 (参见图 4.5.11).

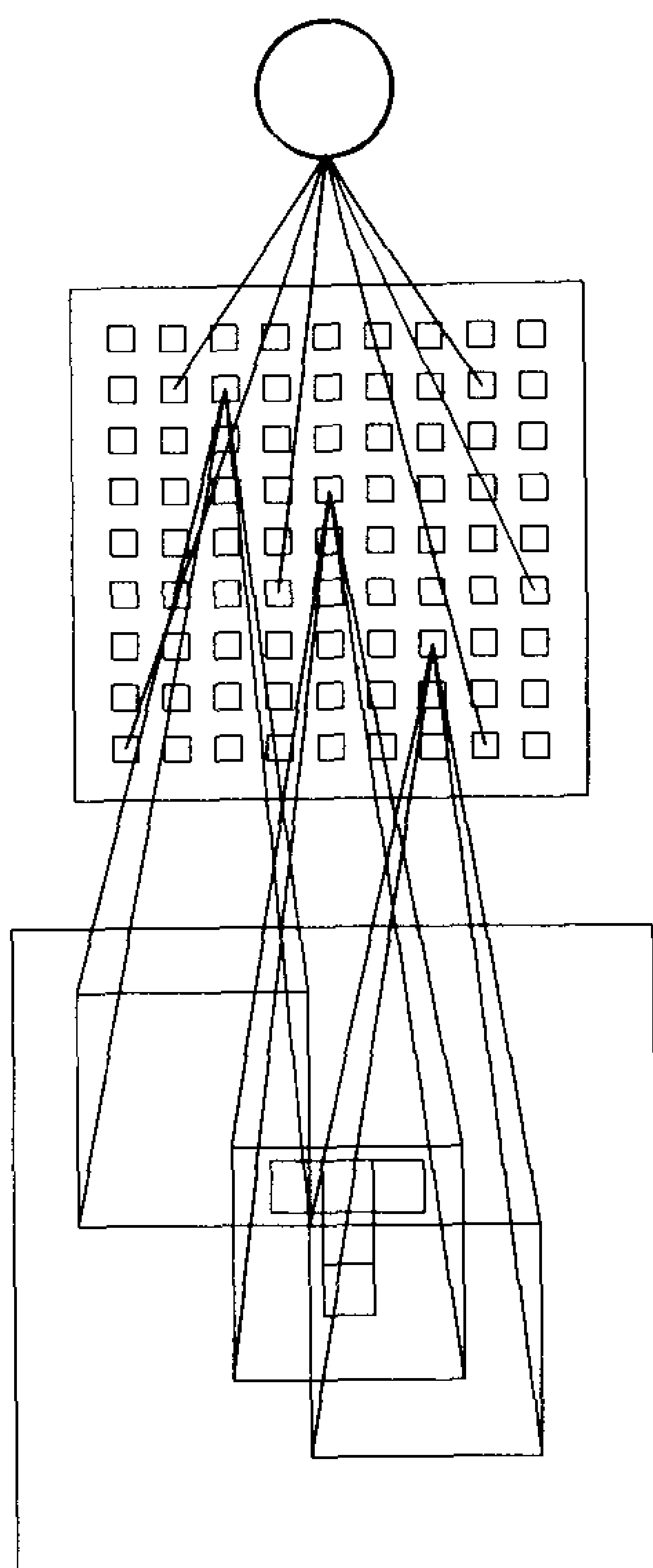


图 4.5.11　T-C 分类的 BP 网络

利用 BP 网络学习算法，求得适当的权和阈值，即能区分输入序列中的字母构成.

图 4.5.12 所示是对于不同的初值用 BP 算法学习后得到的 4 种不同的权.

在图 4.5.12(a) 中，权系数矩阵为

$$w=\begin{pmatrix} -1 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & -1 \end{pmatrix}. \tag{4.5.20}$$

此时对字母 "T"，总存在一个隐单元，其输入加权和为 1 (如图

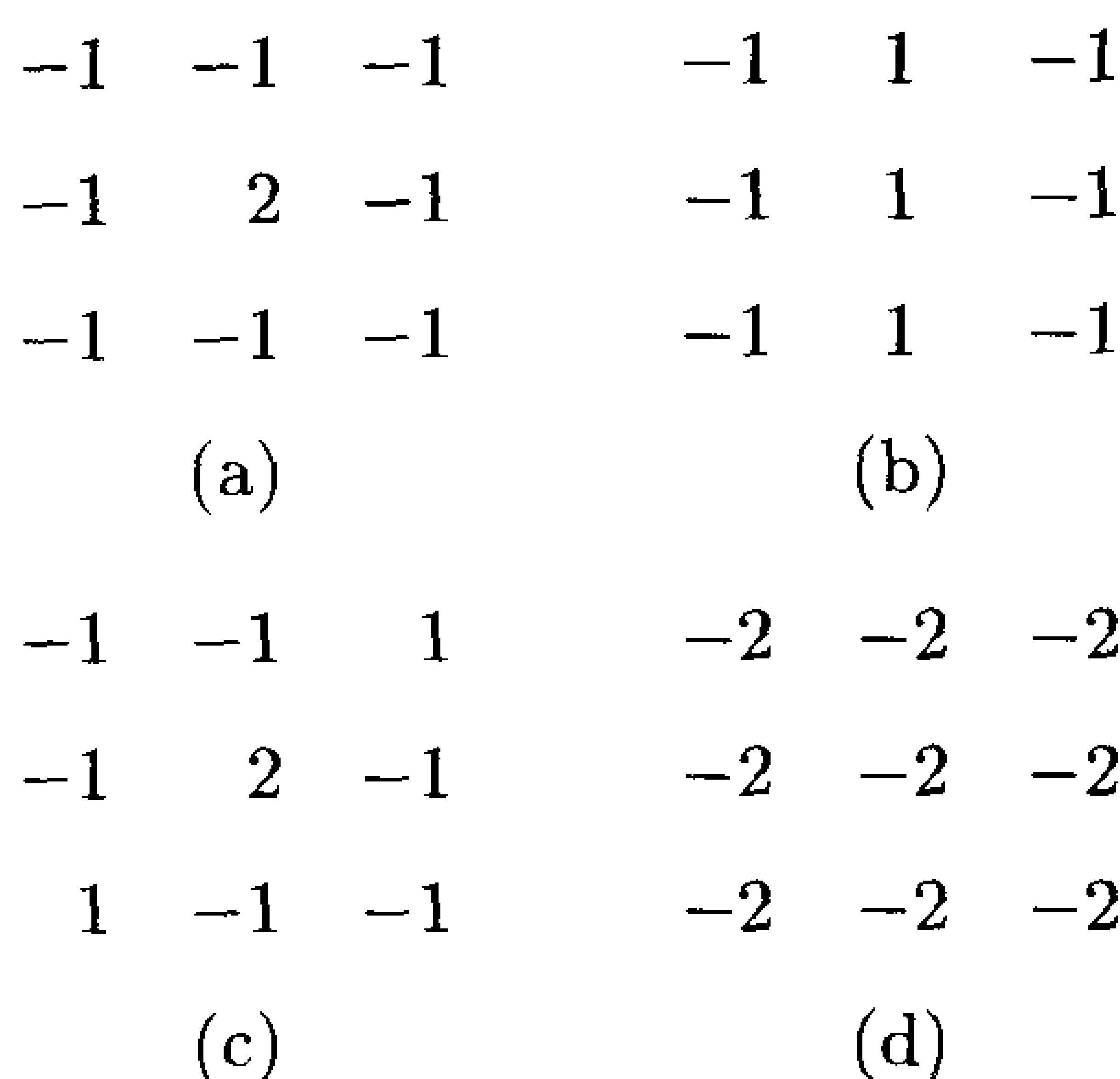


图 4.5.12　用 BP 算法得到的不同类型的权

4.5.13 所示). 由于权系数矩阵是中心对称的, 而字母 "T" 的不同旋转方向构成的输入矩阵也是中心对称的, 因此不论字母 "T" 放在什么方向上, 总存在一个隐单元, 其输入加权和为 1. 而对字母 "C", 不论沿什么方向放置, 其所有隐单元的输入加权和均不大于 0. 故只要取输出单元的阈值 $\theta = 0.5$, 即可将两种字母区分开来, 当输入字母 "T" 时, 输出单元的值为 1, 而当输入字母 "C" 时, 输出单元的值为 0.

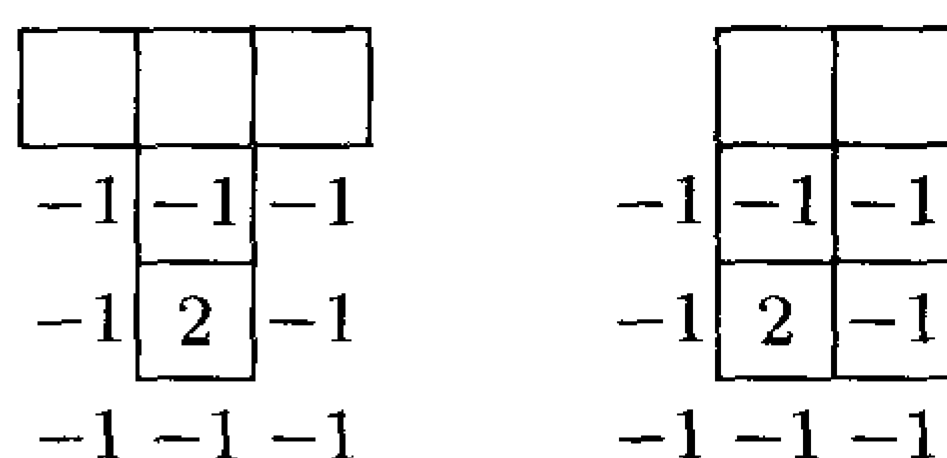


图 4.5.13　在 图 4.5.12(a) 权下, "T" 和 "C" 的输入加权和

同样地, 在图 4.5.12(b) 中, 权系数矩阵为

$$w = \begin{pmatrix} -1 & 1 & -1 \\ -1 & 1 & -1 \\ -1 & 1 & -1 \end{pmatrix}. \tag{4.5.21}$$

若输入字母 “T”, 不论它沿哪一个方向放置，都存在一个隐单元的加权和为 2, 而输入字母 “C”, 则最多有一个隐单元的加权和为 1, 其余均小于 1 (见图 4.5.14). 只要取输出单元的阈值 $\theta = 1.5$, 即可区分两种字母.

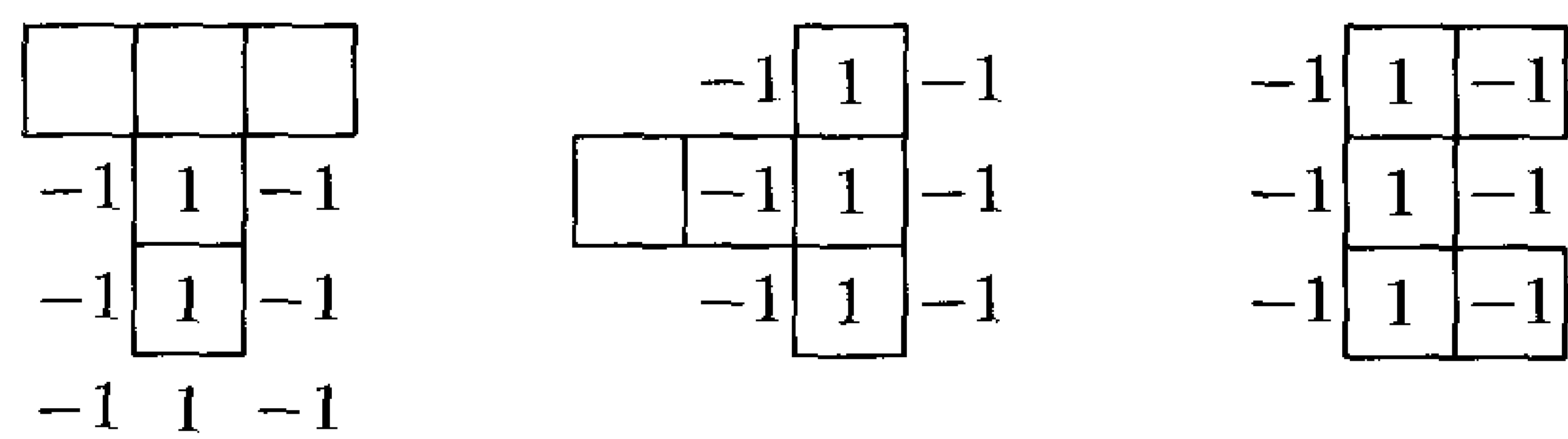


图 4.5.14　在 图 4.5.12(b) 权下，“T” 和 “C” 的输入加权和

在图 4.5.12(c) 中，权系数矩阵为

$$w = \begin{pmatrix} -1 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & -1 \end{pmatrix}. \tag{4.5.22}$$

当输入字母 “T” 时，不论什么方向，都有 5 个隐单元的加权和大于 0, 而当输入字母 “C” 时，只有 3 个隐单元的加权和大于 0 (见图 4.5.15). 取阈值 $\theta = 4$, 即可将字母 “T” 和 “C” 区分开来.

在图 4.5.12(d) 中，权系数矩阵为

$$w = \begin{pmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{pmatrix}. \tag{4.5.23}$$

若输入 “T”, 有 4 个隐单元的加权和非负，而若输入 “C”, 有 5 个隐单元的加权和非负 (见图 4.5.16). 取阈值 $\theta = 4.5$, 即可区分字母 “T” 和 “C”.

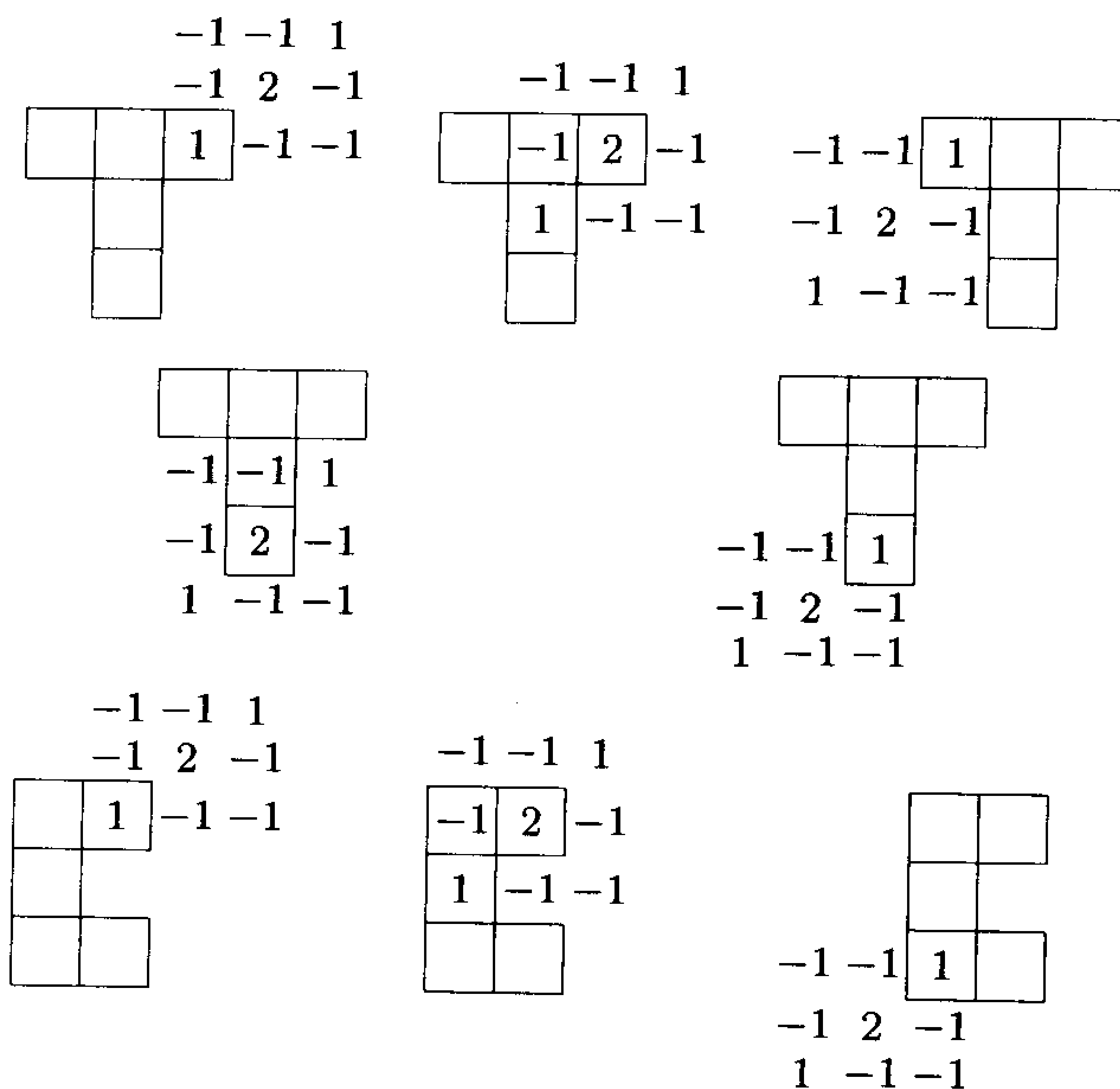


图 4.5.15　在图 4.5.12(c) 权下，“T”和“C”的输入加权和

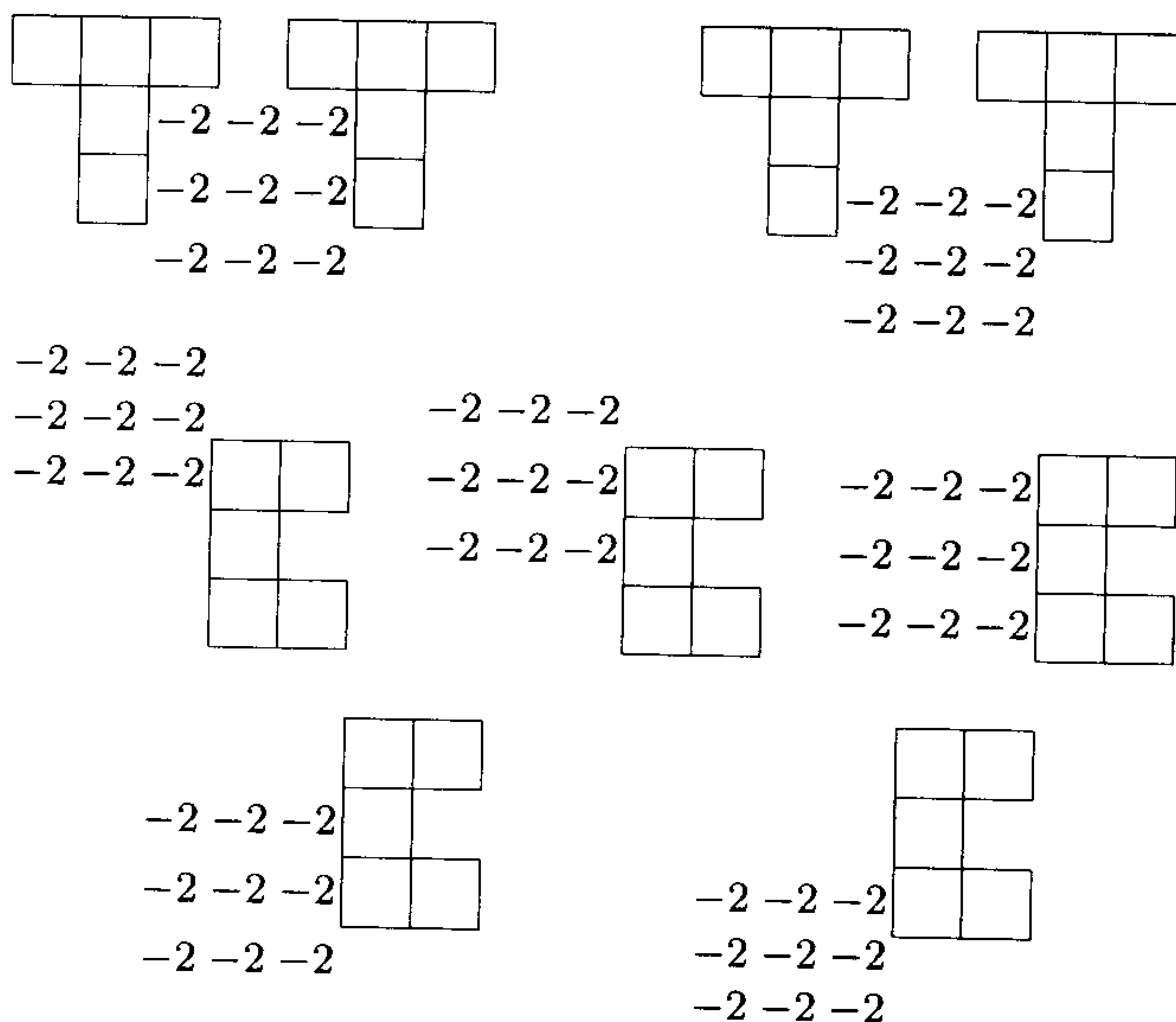


图 4.5.16　在 图 4.5.12(d) 权下，“T”和“C”的输入加权和

§4.6 模 式 识 别

4.6.1 字母识别

人的视网膜可以接受外界的信息，如文字、图形、实物等，并能识别这些信息. 当人看到写在纸上的文字时，不论这个文字放在什么位置，沿什么方向书写，均可以准确地辨别出这些文字. 如果希望用计算机来识别输入的文字图像，就需要有相应的判别算法. 这里我们借助多层自适应神经网络 (Madaline) 来实现字母的识别问题.

多层自适应神经网络 (Madaline) 由多个单层自适应线性神经网络 (Adaline) 组合而成 (参见图 4.6.1，图中的 AD 表示 Adaline).

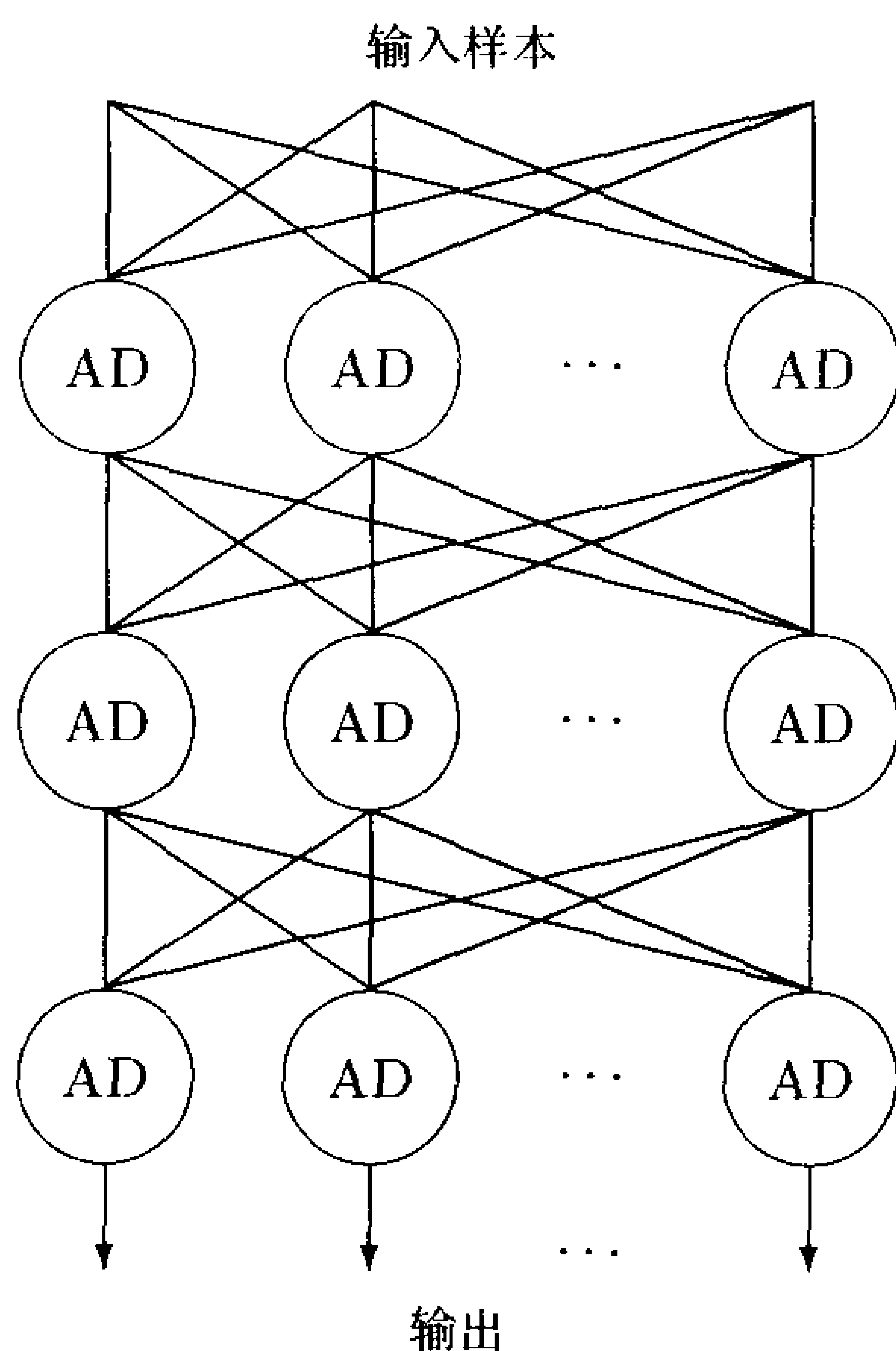


图 4.6.1 多层自适应神经网络的示意图

Madaline 的学习分两步进行：第一步利用 MRII 算法计算出每一层的要求响应，第二步采用单层自适应线性神经网络的 LMS 算法进行学习，得到每一层间的权系数.

设网络的输入为 $x_0=(x_{01},x_{02},\cdots,x_{0n})^{\mathrm{T}}$, 每一层的输出分两个部分：模拟输出部分 y_k 和数字输出部分 x_k

$$y_k=\sum w_{ki}x_i-\theta_k, \tag{4.6.1}$$

$$x_k=\operatorname{sgn}(y_k), \tag{4.6.2}$$

其中数字输出 x_k 又作为下一层的输入. MRII 算法的基本思想是：改变对网络干扰最小的输出量的符号，观察总的网络的输出响应和要求的值之间的误差. 如果误差变小，就接受这种改变，否则恢复原来的符号. 具体的学习步骤如下：

(1) 初始化 Madaline 网络中的权，用一个随机数作为各个权的初始权值;

(2) 输入一个样本向量 x_k 和要求输出 t_k, 由 (4.6.1)~(4.6.2) 式按层一步步地计算出实际输出，并求出要求的输出 t_k 和实际输出的误差;

(3) 根据最小干扰的原则，从第一层开始，寻找第一层中神经元模拟输出最接近于 0 的那个神经元，改变其输出数字量的符号;

(4) 计算网络最后的输出和要求响应之间的误差，如果误差减小，则接受这种改变，否则，恢复原来输出数字量的符号;

(5) 第一个神经元训练结束后，接着训练下一个神经元，即模拟输出第二个接近 0 的单元，仍按上面的规则训练;

(6) 当这一层单个神经元训练完了，再按两个一组进行训练，然后再按 3 个一组进行训练，$\cdots\cdots$ 直到输出与要求响应之间的误差减到最小为止. 然后用同样的方法训练第二层、第三层，$\cdots\cdots$;

(7) 输入另一个样本向量，用同样的方法进行训练，直到误差达到最小为止.

通过 MRII 的学习可以得到每个输入样本对应的每一层相应的输出，从而对每一单层的 Adaline 可以用 LMS 算法进行学习，

得到全部权的解. LMS 在 §4.1.2 中已有详细介绍.

下面我们介绍 Madaline 在字母识别中的应用. 设输入的文字是二维的, 每个文字用一个矩阵 $x = (x_{ij})$ 作为输入单元, 表示点阵文字, 其中的元素为二值量, 当该点为白色时, $x_{ij} = -1$; 当该点是黑色时, $x_{ij} = 1$.

为了使输入的文字样本放在任何位置、沿任何方向都能识别, 就要求由多层自适应线性神经单元组成的网络是不变性网络, 它的输出只与不同输入模式有关, 而与同一输入模式的位置、方向无关.

图 4.6.2 是字母识别的 Madaline 的示意图.

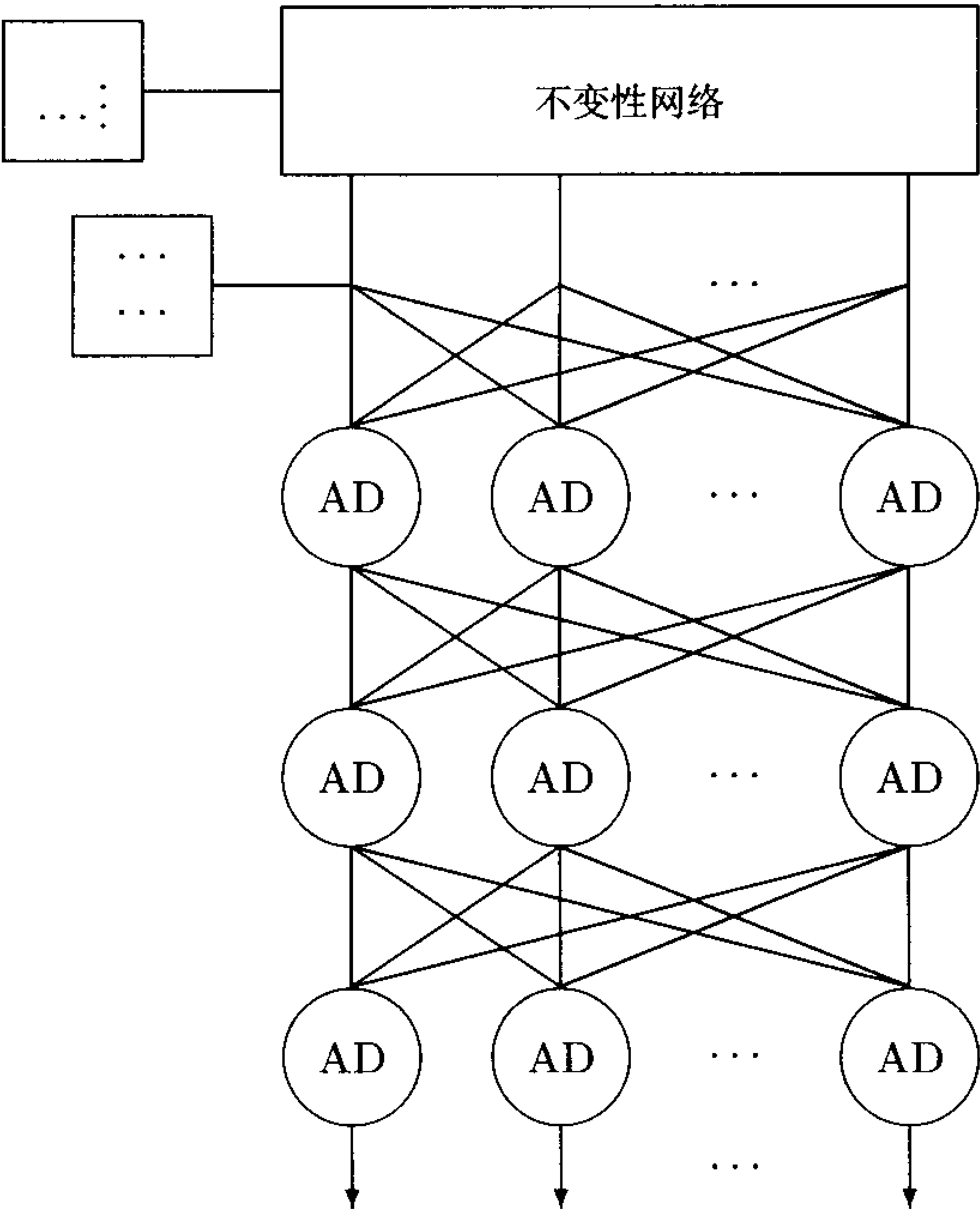


图 4.6.2 字母识别的 Madaline 的示意图

考虑到网络是旋转不变的, 因此几个子平面的组合, 对应于一个输出, 从而不变性网络包含很多子平面, 并组合为一定的输

出，每个子平面如图 4.6.3 所示，图中的 AD 表示 Adaline. 每一个子平面上有很多 AD 按方阵排列，输入视网膜上所有的元都与每个 AD 相联，而每个 AD 的输出都连接到一个选择器 MAJ 上. 如果在子平面左上角的 Adaline 与输入二维图像的连接权矩阵为 w，那么子平面上其它 AD 与输入单元的连接权分布如下：

$$\begin{pmatrix} w & R_1(w) & R_2(w) & R_3(w) \\ D_1(w) & D_1R_1(w) & D_1R_2(w) & D_1R_3(w) \\ D_2(w) & D_2R_1(w) & D_2R_2(w) & D_2R_3(w) \\ D_3(w) & D_3R_1(w) & D_3R_2(w) & D_3R_3(w) \end{pmatrix}, \qquad (4.6.3)$$

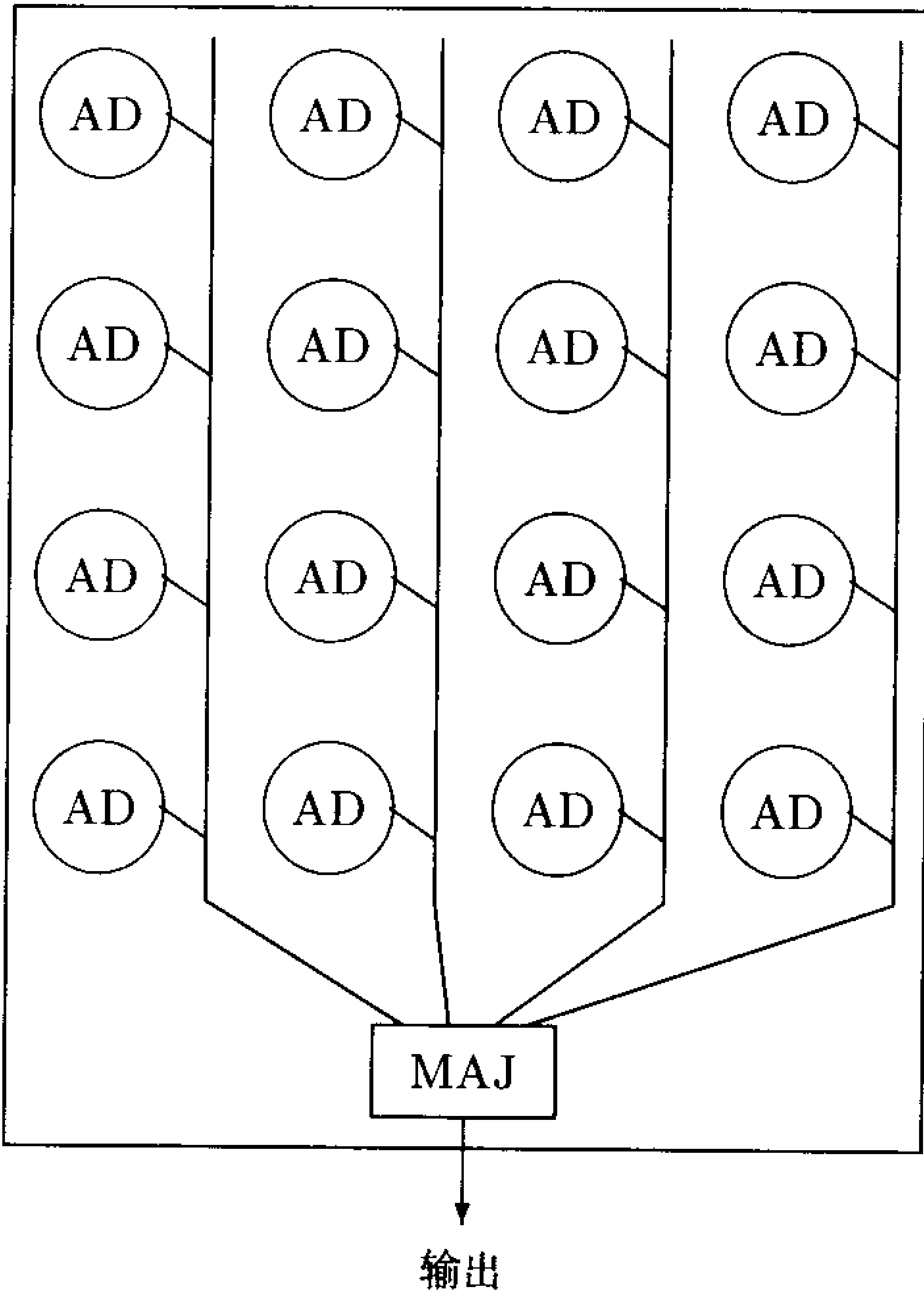


图 4.6.3　不变性网络的子平面结构示意图

其中 R_1 表示权矩阵 w 向右循环移一列，D_1 表示权矩阵 w 向下循环移一行，$D_1R_1(w)$ 表示权矩阵 w 向右循环移一列并向下循环

移一行，即

$$w = \begin{pmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{pmatrix}, \tag{4.6.4}$$

$$R_1(w) = \begin{pmatrix} w_{14} & w_{11} & w_{12} & w_{13} \\ w_{24} & w_{21} & w_{22} & w_{23} \\ w_{34} & w_{31} & w_{32} & w_{33} \\ w_{44} & w_{41} & w_{42} & w_{43} \end{pmatrix}, \tag{4.6.5}$$

$$D_1(w) = \begin{pmatrix} w_{41} & w_{42} & w_{43} & w_{44} \\ w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \end{pmatrix}, \tag{4.6.6}$$

$$D_1R_1(w) = \begin{pmatrix} w_{44} & w_{41} & w_{42} & w_{43} \\ w_{14} & w_{11} & w_{12} & w_{13} \\ w_{24} & w_{21} & w_{22} & w_{23} \\ w_{34} & w_{31} & w_{32} & w_{33} \end{pmatrix}. \tag{4.6.7}$$

w 的权是随机产生的，一旦产生了，其它 AD 的权都按照上述的方法向下或向右循环移动 k 次， k 为 $1,2,\cdots,n$. 因此这样的结构，不管输入样本上下左右发生什么变化，在子平面上 Adaline 输出的 $+1$ 或 -1 的数目是不变的. 选择器 MAJ 把它们的输出数值累加起来经一定的组合输出，它对平移是不变的. 对于旋转不变可用下面 4 个子平面的综合获得，如图 4.6.4 表示输入样本在旋转 $90°,180°,270°$ 时的情况. 设 s_1 平面是最左面的平面，表示没有旋转，其连接权矩阵由 (4.6.1) 式给出，那么 s_2 为旋转 $90°$ 时的情况，其权的分布如下

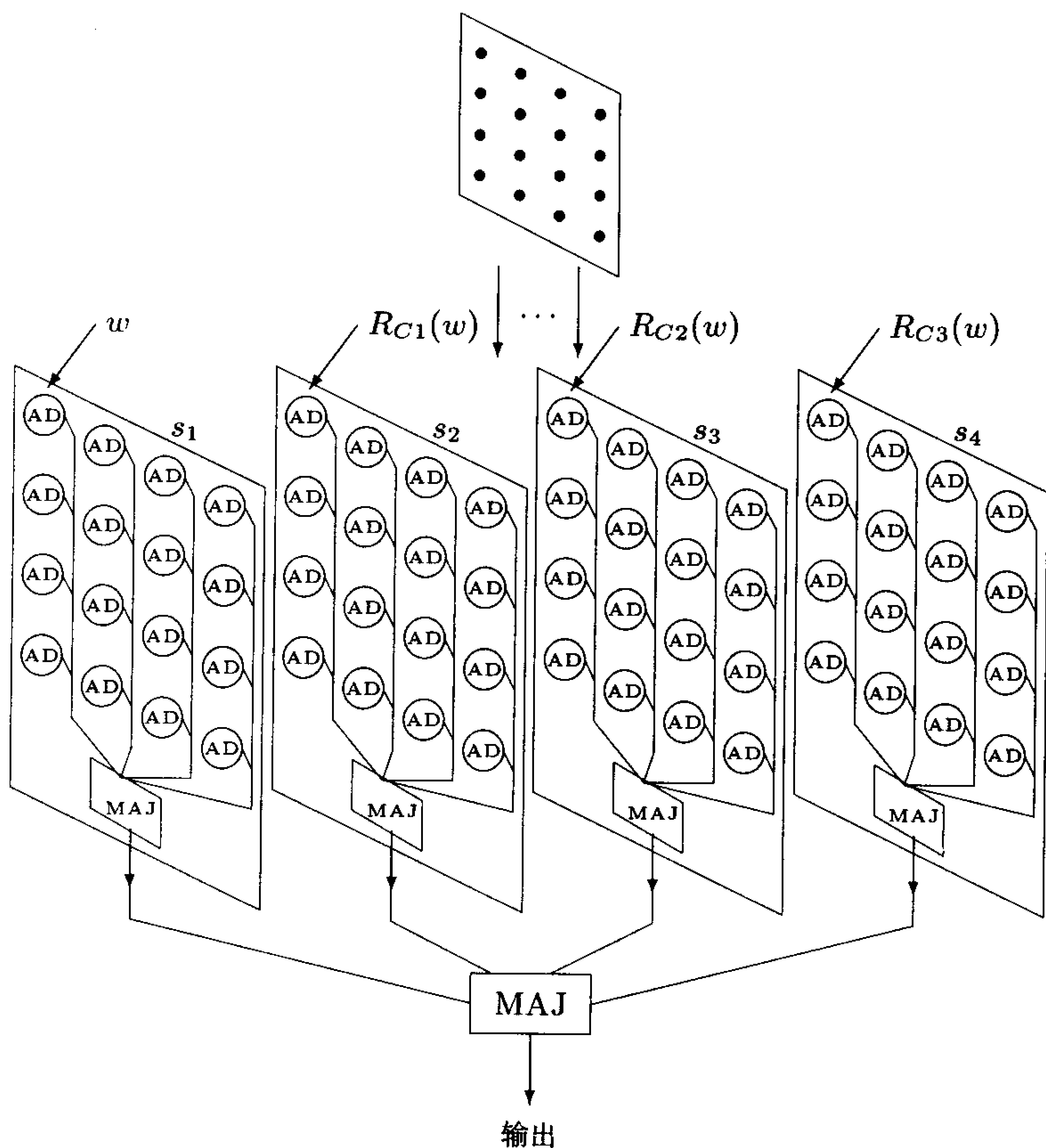


图 4.6.4　旋转不变的网络结构示意图

$$\begin{pmatrix} R_{C1}(w) & R_1R_{C1}(w) & R_2R_{C1}(w) & R_3R_{C1}(w) \\ D_1R_{C1}(w) & D_1R_1R_{C1}(w) & D_1R_2R_{C1}(w) & D_1R_3R_{C1}(w) \\ D_2R_{C1}(w) & D_2R_1R_{C1}(w) & D_2R_2R_{C1}(w) & D_2R_3R_{C1}(w) \\ D_3R_{C1}(w) & D_3R_1R_{C1}(w) & D_3R_2R_{C1}(w) & D_3R_3R_{C1}(w) \end{pmatrix}, \tag{4.6.8}$$

其中 $R_{C1}(w)$ 表示 w 矩阵旋转 90° 的权矩阵. s_3, s_4 分别是旋转 180°, 270° 时的平面. 它们的输出又综合到一个总的 MAJ 输出.

这样对应于一个样本, 可能需要有 4 个 AD 子平面来保证其

平移和旋转不变. 对于 k 个样本, 最小需要的子平面数为 $4\log_2 k$, 其中 "4" 为旋转的 4 个方向, $\log_2 k$ 为能与样本一一对应的最小编码.

4.6.2 手写数字的模式识别

随着计算机的飞速发展, 文字的输入方法也有了很大的改进. 目前已有一些特制的笔 (如汉王笔), 用户使用这种笔即可像平常写字一样, 通过手写文字就可以达到输入的目的. 要实现手写输入的目的, 需要对手写的文字加以学习识别, 辨认出不同用户各种不同的字体. 由于用户手写文字的差别很大, 用传统的模式识别的方法很难得到高的识别率, 需要提供一种识别率更高的算法.

这里利用 BP 网络学习方法, 对手写数字进行识别. BP 网络综合了传统的统计模式识别和句法模式识别的特点, 它的权是通过对训练样本的学习得到的, 因而有较高的识别率. 由于不同人的手写数字, 其体形大小不尽相同, 这里取其特征作为 BP 网络的输入. 例如考虑端点特征, 1, 2, 3, 7 都有两个端点, 0, 6, 8, 9 都是由圈构成的; 在 1, 2, 3, 7 这组数字中, 2 的两个端点为上左下右, 3 的两个端点为两个都在一边, 而在 0, 6, 8, 9 这组数字中, 9 的圈在上, 6 的圈在下. 每个特征在输入神经元上取 1, 而无此特征则取 0, 这里共取了 34 个特征作为区分不同字型的依据. 当然随着学习样本数的增加, 可能要增加输入的特征数. 如果输出神经元 $m = 10$, 对应 10 个数字, 其输出只有一个为 1, 其它为 0, 这样的分类输入和输出都是 0 和 1 两个值, 中间层是 $[0,1]$ 之间的模拟量, 因此, 可用 3 层 BP 网络来完成, 其隐单元数可由以下 3 个公式之一给出

$$k < \sum_{i=0}^{n} C_{n_1}^{i}, \tag{4.6.9}$$

$$n_1 = \sqrt{n+m} + a, \tag{4.6.10}$$

$$n_1 = \log_2 n, \tag{4.6.11}$$

其中 k 为样本数, n_1 为隐单元数, n 为输入单元数, m 为输出

单元数， a 为 $1\sim10$ 之间的正整数.

对于 34 个输入单元 (对应于 34 个特征) 可能的输入样本多达 2^{34} 个，但是实际上很多样本是不会出现的. 根据 200 人写的数字统计，样本数约为 $k=1000$. 由 (4.6.9) 式，取 $n_1=10$, 有

$$\sum_{i=0}^{n} C_{n_1}^{i} = C_{10}^{0} + C_{10}^{1} + \cdots + C_{10}^{10} = 1024 > 1000. \tag{4.6.12}$$

由 (4.6.10) 式,

$$n_1 = \sqrt{m+n} + a = \sqrt{44} + a = 8 \sim 17, \tag{4.6.13}$$

其中 $a=1\sim10$.

由 (4.6.11) 式,

$$n_1 = \log_2 n = 6. \tag{4.6.14}$$

由 (4.6.12)~(4.6.14) 式，并考虑到对测试样本的容差性，取 $n_1=14$. 通过对样本的学习，并对 6000 多个手写数字的测试，其识别率大于 95%.

4.6.3 运动物体的速度场计算

物体运动时，从两幅不同时间拍摄的照片上计算出物体上每个像素的速度方向和大小, 以表示物体的运动情况, 这种像素速度方向的组合形成的图形称为 **速度场**. 由于照片是二维的，这里主要考虑二维图像速度场的计算, 它的条件为视角小, 使物体没有偏离相机的视平面. 假设物体是刚性的, 物体的光强在运动时不变, 那么运动物体的速度场可以采用神经网络来计算.

由于物体是刚性的，因而速度场是连续的，即如果图形的 x, y 方向的速度分别用

$$u = \frac{\mathrm{d}x}{\mathrm{d}t}, \quad v = \frac{\mathrm{d}y}{\mathrm{d}t} \tag{4.6.15}$$

来表示，那么输出速度场的速度在二维平面上是一个连续函数. 考虑目标函数

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \tag{4.6.16}$$

达到最小，即速度的变化率应该最小.

如果在一张照片的二维图像上每个点的光强为 $I(x,y)$, 那么总光强为

$$I=\sum_{x,y} I(x,y).$$

由于总光强也是不变的，因而有

$$\frac{\mathrm{d}I}{\mathrm{d}t}=0,$$

即

$$\sum_{x,y}\left(\frac{\partial I}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{\partial I}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}+\frac{\mathrm{d}I}{\mathrm{d}t}\right)=0.$$

记

$$\frac{\partial I}{\partial x}=I_x,\quad \frac{\partial I}{\partial y}=I_y,\quad \frac{\mathrm{d}I}{\mathrm{d}t}=I,$$

得到能量函数为

$$\begin{aligned}E(u,v)=\iint\Bigg\{(I_xu+I_yv+I_t)^2+\lambda\Bigg[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2\\+\left(\frac{\partial v}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2\Bigg]\Bigg\}\mathrm{d}x\mathrm{d}y,\end{aligned}\tag{4.6.17}$$

其中 λ 是一个常数.

对能量函数进行变分运算，在稳定时 $\delta E=0$, 得到两个线性方程

$$I_x^2u+I_xI_yv+I_xI_t-\lambda\nabla^2u=0,\tag{4.6.18}$$

$$I_xI_yu+I_y^2v+I_yI_t-\lambda\nabla^2v=0.\tag{4.6.19}$$

将上面两个线性方程写成离散形式. 把图像划分成许多小的格点，它们均匀分布在二维空间内. 若用 I_{xij}, I_{yij}, I_{tij} 分别表示在某一格点 (i,j) 上的光强随 x, y, t 的变化率，用 u_{ij}, v_{ij} 表示格

点 (i,j) 的 x, y 方向的速度，而 $\nabla^2 u, \nabla^2 v$ 用二阶差商表示：

$$\begin{aligned}\nabla^2 u &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \\ &= (u_{i,j-1} + u_{i,j+1} + u_{i+1,j} + u_{i-1,j}) - 4u_{ij},\end{aligned} \tag{4.6.20}$$

$$\begin{aligned}\nabla^2 v &= \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \\ &= (v_{i,j-1} + v_{i,j+1} + v_{i+1,j} + v_{i-1,j}) - 4v_{ij},\end{aligned} \tag{4.6.21}$$

那么它与周围格点之间的关系可写成下列形式：

$$\begin{aligned}&I_{xij}^2 u_{ij} + I_{xij} I_{yij} v_{ij} + I_{xij} I_{tij} \\ &-\lambda(u_{i+1,j} + u_{i,j+1} + u_{i-1,j} + u_{i,j-1} - 4u_{ij}) = 0,\end{aligned} \tag{4.6.22}$$

$$\begin{aligned}&I_{xij} I_{yij} u_{ij} + I_{yij}^2 v_{ij} + I_{yij} I_{tij} \\ &-\lambda(v_{i+1,j} + v_{i,j+1} + v_{i-1,j} + v_{i,j-1} - 4v_{ij}) = 0.\end{aligned} \tag{4.6.23}$$

而利用 Hopfield 神经网络算法，将 (4.6.17) 式定义的函数 E 视为 Hopfield 能量函数，u_j 视为状态变量，显然 $E > 0$. 取状态变量 u_{ij} 满足状态方程

$$\frac{\partial E}{\partial u_{ij}} = -\frac{\mathrm{d}u_{ij}}{\mathrm{d}t},$$

可得

$$\begin{aligned}C\frac{\mathrm{d}u_{ij}}{\mathrm{d}t} &= T(u_{i+1,j} + u_{i,j+1} + u_{i-1,j} + u_{i,j-1} - 4u_{ij}) \\ &\quad + g_{ij}^u(E_{ij} - u_{ij}) + T_{c-i,j}(v_{ij} - u_{ij}),\end{aligned} \tag{4.6.24}$$

$$\begin{aligned}C\frac{\mathrm{d}v_{ij}}{\mathrm{d}t} &= T(v_{i+1,j} + v_{i,j+1} + v_{i-1,j} + v_{i,j-1} - 4v_{ij}) \\ &\quad + g_{ij}^v(E_{ij} - v_{ij}) + T_{c-i,j}(u_{ij} - v_{ij}),\end{aligned} \tag{4.6.25}$$

其中

$$T = \lambda,$$

$$g_{ij}^u = I_{xij}(I_{xij} + I_{yij}),$$

$$g_{ij}^{v} = I_{yij}(I_{xij} + I_{yij}),$$

$$E_{ij} = \frac{-I_{tij}}{I_{xij} + I_{yij}},$$

$$T_{c-i,j} = -I_{xij}I_{yij}.$$

比较 (4.6.22) 和 (4.6.24) 式, (4.6.23) 和 (4.6.25) 式, 显然它们是一致的. 而 E_{ij}, g_{ij}^{u}, g_{ij}^{v} 都是与相片上的光强有关的量, 因此要完成 (4.6.24)~(4.6.25) 的网络可采用光敏元件和电阻网络相结合的电路. 图 4.6.5 就是完成状态方程 (4.6.24) 的一个电阻网络, 在格点 (i,j) 上输入 E_{ij}, 接入导纳 g_{ij}^{u} 和电容 C, 并且电阻 T 与其它格点相连. 总的网络形状如图 4.6.6 所示, 它是由两层如图 4.6.4

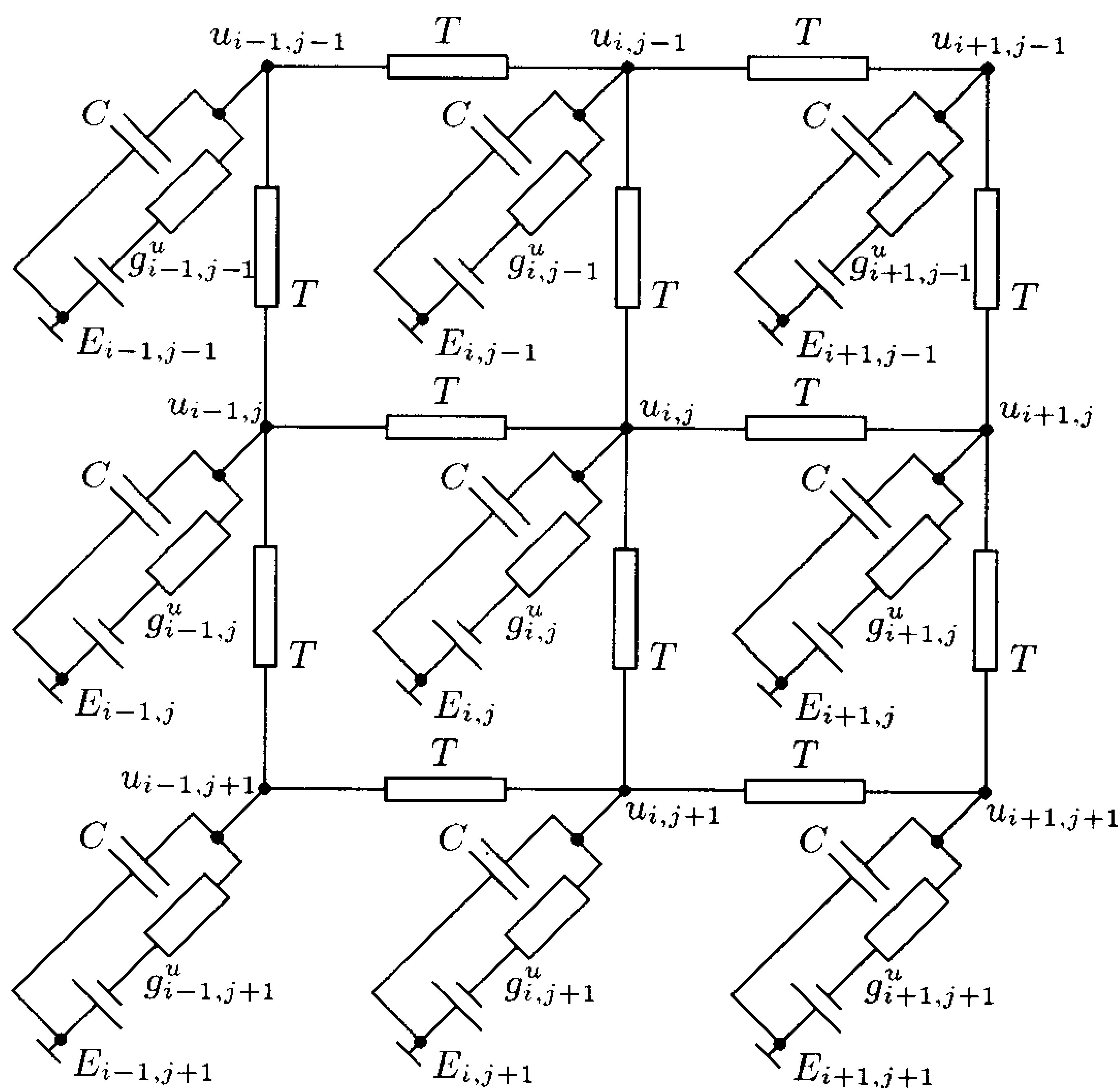


图 4.6.5 完成状态方程 (4.6.24) 的电阻网络

的网络组成，都与光敏器件相连，两层之间又用 $T_{c-i,j}$ 相连接，当运动物体在光敏器件发出 E_{ij}, g_{ij}^u, g_{ij}^v; 加入到网络，在网络的各格点 (i,j) 上输出为 u_{ij}, v_{ij} 的值.

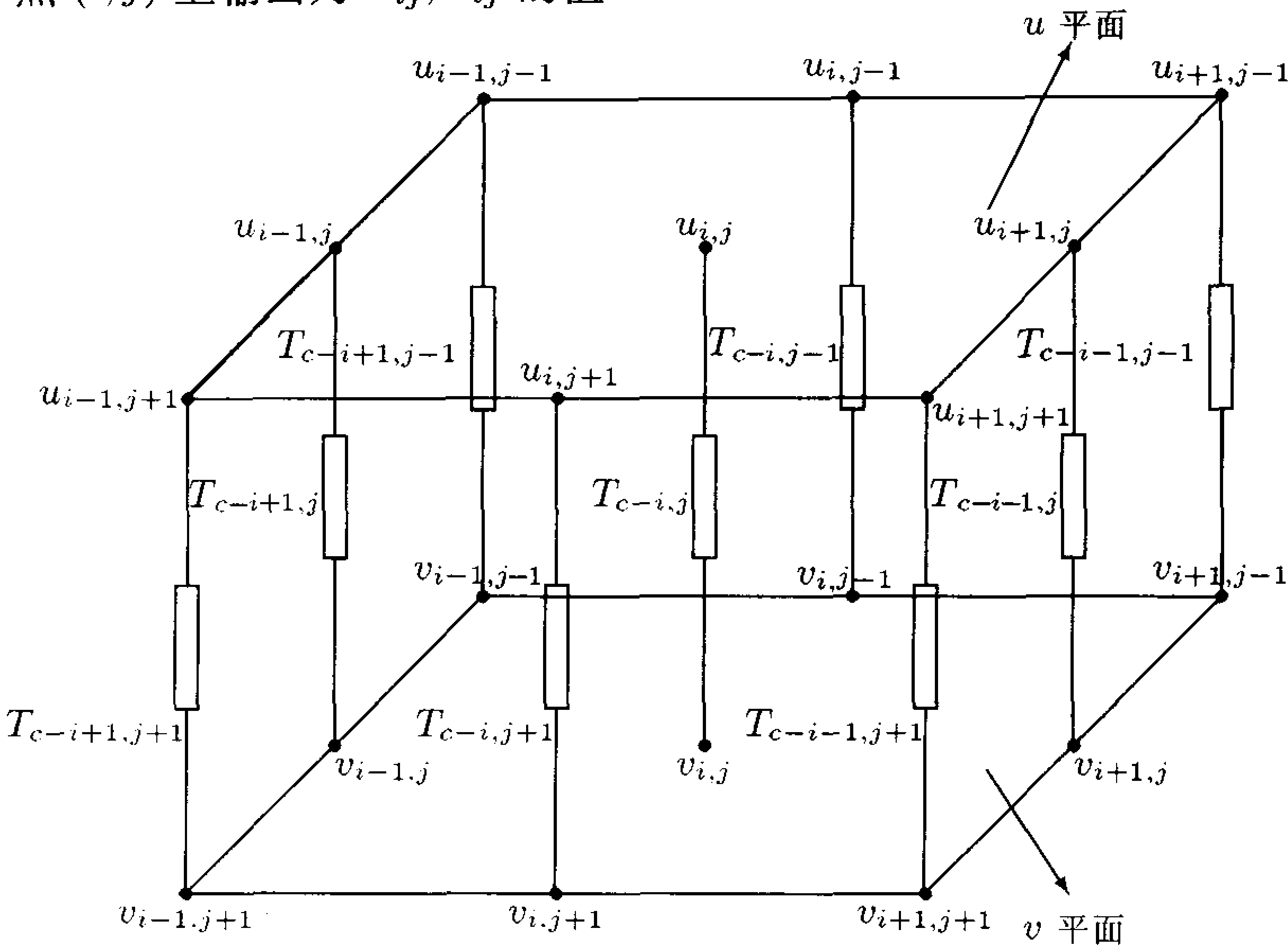


图 4.6.6　速度场计算的电阻网络

参 考 文 献

[1] Cai Zhijie, Ruan Jiong & Li Lijuan, An algorithm of predictability and its application, Communications in Nonlinear Science & Numberical Simulation, 4:1(1999), 54~56.

Proceedings of Satellite Symposium of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Technology Exchange Ltd., Beijing, 1998, 145~146.

[2] D. M. Wolpert & R. C. Miall, Detecting chaos with neural networks, Proc. R. Soc. Lond. B, 242(1990), 82~86.

[3] A. S. Miller, B. H. Blott & T. K. Hames, Review of neural network

applications in medical imaging and signal processing, Med. & Biol. Eng. & Comput., 30(1992), 449~464.

[4] B. Widrow & R. Winter, Neural nets for adaptive filtering and adaptive pattern recognition, IEEE Computer, (1988), 25~39.

[5] 顾凡及等, 一种修正的动态神经元网络模型的类脑电活动及其应用, 生物物理学报, 第 11 卷第 4 期 (1995), 551~556.

[6] 许世雄、阮炯、蔡志杰、蒋云峰, 股市扩容的资金流平衡流体力学模型研究, 中国工业与应用数学学会第四次大会论文集 (曾庆存、李大潜主编), 复旦大学出版社, 1996, 364~368.

[7] 焦李成, 神经网络系统理论, 西安电子科技大学出版社, 1990.

[8] 张立明, 人工神经网络的模型及其应用, 复旦大学出版社, 1993.